<u>Isoform 1:</u>

```
   1 TTGCTCACTG CTCACCCACC TGCTGCTGCC ATGAGGCACC TTGGGGCCTT
  51 CCTCTTCCTT CTGGGGGTCC TGGGGGCCCT CACTGAGATG TGTGAAATAC
 101 CAGAGATGGA CAGCCATCTG GTAGAGAAGT TGGGCCAGCA CCTCTTACCT
 151 TGGATGGACC GGCTTTCCCT GGAGCACTTG AACCCCAGCA TCTATGTGGG
 201 CCTACGCCTC TCCAGTCTGC AGGCTGGGAC CAAGGAAGAC CTCTACCTGC
 251 ACAGCCTCAA GCTTGGTTAC CAGCAGTGCC TCCTAGGGTC TGCCTTCAGC
 301 GAGGATGACG TGACTGCCA GGGCAAGCCT TCCATGGGCC AGCTGGCCCT
 351 CTACCTGCTC GCTCTCAGAG CCAACTGTGA GTTTGTCAGG GGCCACAAGG
 401 GGGACAGGCT GGTCTCACAG CTCAAATGGT TCCTGGAGGA TGAGAAGAGA
 451 GCCATTGACA CAGCAGCCAT GGCAGGCTTG GCATTCACCT GTCTGAAGCG
 501 CTCAAACTTC AACCCTGGTC GGAGACAACG GATCACCATG GCCATCAGAA
 551 CAGTGCGAGA GGAGATCTTG AAGGCCCAGA CCCCCGAGGG CCACTTTGGG
 601 AATGTCTACA GCACCCCATT GGCATTACAG TTCCTCATGA CTTCCCCCAT
 651 GCGTGGGGCA GAACTGGGAA CAGCATGTCT CAAGGCGAGG GTTGCTTTGC
 701 TGGCCAGTCT GCAGGATGGA GCCTTCCAGA ATGCTCTCAT GATTTCCCAG
 751 CTGCTGCCCG TTCTGAACCA CAAGACCTAC ATTGATCTGA TCTTCCCAGA
 801 CTGTCTGGCA CCACGAGTCA TGTTGGAACC AGCTGCTGAG ACCATTCCTC
 851 AGACCCAAGA GATCATCAGT GTCACGCTGC AGGTGCTTAG TCTCTTGCCG
 901 CCGTACAGAC AGTCCATCTC TGTTCTGGCC GGGTCCACCG TGGAAGATGT
 951 CCTGAAGAAG GCCCATGAGT TAGGAGGATT CACATATGAA ACACAGGCCT
1001 CCTTGTCAGG CCCCTACTTA ACCTCCGTGA TGGGGAAAGC GGCCGGAGAA
1051 AGGGAGTTCT GGCAGCTTCT CCGAGACCCC AACACCCCAC TGTTGCAAGG
1101 TATTGCTGAC TACAGACCCA AGGATGGAGA AACCATTGAG CTGAGGCTGG
1151 TTAGCTGGTA GCCCCTGAGC TCCCTCATCC CAGCAGCCTC GCACACTCCC
1201 TAGGCTTCTA CCCTCCCTCC TGATGTCCCT GGAACAGGAA CTCGCCTGAC
1251 CCTGCTGCCA CCTCCTGTGC ACTTTGAGCA ATGCCCCCTG GGATCACCCC
1301 AGCCACAAGC CCTTCGAGGG CCCTATACCA TGGCCCACCT TGGAGCAGAG
1351 AGCCAAGCAT CTTCCCTGGG AAGTCTTTCT GGCCAAGTCT GGCCAGCCTG
1401 GCCCTGCAGG TCTCCCATGA AGGCCACCCC ATGGTCTGAT GGGCATGAAG
1451 CATCTCAGAC TCCTTGGCAA AAAACGGAGT CCGCAGGCCG CAGGTGTTGT
1501 GAAGACCACT CGTTCTGTGG TTGGGGTCCT GCAAGAAGGC CTCCTCAGCC
1551 CGGGGGCTAT GGCCCTGACC CCAGCTCTCC ACTCTGCTGT TAGAGTGGCA
1601 GCTCCGAGCT GGTTGTGGCA CAGTAGCTGG GGAGACCTCA GCAGGGCTGC
1651 TCAGTGCCTG CCTCTGACAA AATTAAAGCA TTGATGGCCT GTGAAAAAAA
1701 AAAAAAAAAA AAAAAAAAAA AA
```
(SEQ.ID NO:1)

FEATURES:
5'UTR: 1 - 30
Start Codon: 31
Stop Codon: 1159
3'UTR: 1162

Homologous proteins:
Top 10 BLAST Hits

	Score	E
CRA|108000024653390 /altid=gi|12742775 /def=ref|XP_009922.2| tr...	752	0.0
CRA|108000024636236 /altid=gi|298316 /def=gb|AAB25526.1| transc...	732	0.0
CRA|18000004926133 /altid=gi|339205 /def=gb|AAA61057.1| (L02648...	732	0.0
CRA|108000024042036 /altid=gi|12654675 /def=gb|AAH01176.1|AAH01...	731	0.0
CRA|18000004926130 /altid=gi|4507409 /def=ref|NP_000346.1| tran...	727	0.0
CRA|18000004926132 /altid=gi|339203 /def=gb|AAA61056.1| (L02647...	725	0.0
CRA|18000005170902 /altid=gi|7657639 /def=ref|NP_056564.1| tran...	515	e-145
CRA|18000005218941 /altid=gi|4572454 /def=gb|AAD23829.1|AF12128...	501	e-140
CRA|164000136745249 /altid=gi|11968124 /def=ref|NP_071979.1| tr...	481	e-134
CRA|18000004926134 /altid=gi|4507407 /def=ref|NP_001053.1| tran...	108	2e-22

FIGURE 1A

```
gi|10725490 /dataset=dbest /taxon=96...              858   0.0
gi|10947399 /dataset=dbest /taxon=96...              846   0.0
gi|9121897 /dataset=dbest /taxon=9606...             846   0.0
gi|13280819 /dataset=dbest /taxon=96...              846   0.0
gi|13287907 /dataset=dbest /taxon=96...              833   0.0
gi|13286505 /dataset=dbest /taxon=96...              831   0.0
gi|8150776 /dataset=dbest /taxon=960...              815   0.0
gi|5936410 /dataset=dbest /taxon=9606 ...            726   0.0
gi|6888875 /dataset=dbest /taxon=9606...             726   0.0
gi|6888872 /dataset=dbest /taxon=9606...             726   0.0
```

EXPRESSION INFORMATION FOR MODULATORY USE:

library source:

```
gi|10725490|adult adrenal gland
gi|10947399| mammary gland
gi|9121897| retinoblastoma
gi|13280819| adenocarcinoma cell line
gi|13287907| retinoblastoma
gi|13286505| embryonal carcinoma, cell line
gi|8150776| adult uterus
gi|5936410| adult uterus
gi|6888875| adult head_neck
gi|6888872| adult head_neck
```

Tissue Expression:

Human leukocyte

FIGURE 1B

<u>Isoform 2:</u>

```
   1 GGAGGATTAA TCAGTGACAG GAAGCTGCGT CTCTCGGAGC GGTGACCAGC
  51 TGTGGTCAGG AGAGCCTCAG CAGGGCCAGC CCCAGGAGTC TTTCCCGATT
 101 CTTGCTCACT GCTCACCCAC CTGCTGCTGC CATGAGGCAC CTTGGGGCCT
 151 TCCTCTTCCT TCTGGGGGTC CTGGGGGCCC TCACTGAGAT GTGTGAAATA
 201 CCAGAGATGG ACAGCCATCT GGTAGAGAAG TTGGGCCAGC ACCTCTTACC
 251 TTGGATGGAC CGGCTTTCCC TGGAGCACTT GAACCCCAGC ATCTATGTGG
 301 GCCTACGCCT CTCCAGTCTG CAGGCTGGGA CCAAGGAAGA CCTCTACCTG
 351 CACAGCCTCA TGCTTGGTTA CCAGCAGTGC CTCCTAGGGT CTGCCTTCAG
 401 CGAGGATGAC GGTGACTGCC AGGGCAAGCC TTCCATGGGC CAGCTGGCCC
 451 TCTACCTGCT CGCTCTCAGA GCCAACTGGC ATGATCACAA GGGCCACCCC
 501 CACACTAGCT ACTACCAGTA TGGCCTGGGC ATTCTGGCCC TGTGTCTCCA
 551 CCAGAAGCGG GTCCATGACA GCGTGGTGGA CAAACTTCTG TATGCTGTGG
 601 AACCTTTCCA CCAGGGCCAC CATTCTGTGG ACACAGCAGC CATGGCAGGC
 651 TTGGCATTCA CCTGTCTGAA GCGCTCAAAC TTCAACCCTG GTCGGAGACA
 701 ACGGATCACC ATGGCCATCA GAACAGTGCG AGAGGAGATC TTGAAGGCCC
 751 AGACCCCCGA GGGCCACTTT GGGAATGTCT ACAGCACCCC ATTGGCATTA
 801 CAGTTCCTCA TGACTTCCCC CATGCGTGGG GCAGAACTGG AACAGCATG
 851 TCTCAAGGCG AGGGTTGCTT TGCTGGCCAG TCTGCAGGAT GGAGCCTTCC
 901 AGAATGCTCT CATGATTTCC CAGCTGCTGC CCGTTCTGAA CCACAAGACC
 951 TACATTGATC TGATCTTCCC AGACTGTCTG GCACCACGAG TCATGTTGGA
1001 ACCAGCTGCT GAGACCATTC CTCAGACCCA AGAGATCATC AGTGTCACGC
1051 TGCAGGTGCT TAGTCTCTTG CCGCCGTACA GACAGTCCAT CTCTGTTCTG
1101 GCCGGGTCCA CCGTGGAAGA TGTCCTGAAG AAGGCCCATG AGTTAGGAGG
1151 ATTCACATAT GAAACACAGG CCTCCTTGTC AGGCCCCTAC TTAACCTCCG
1201 TGATGGGGAA AGCGGCCGGA GAAAGGGAGT TCTGGCAGCT TCTCCGAGAC
1251 CCCAACACCC CACTGTTGCA AGGTATTGCT GACTACAGAC CCAAGGATGG
1301 AGAAACCATT GAGCTGAGGC TGGTTAGCTG GTAGCCCCTG AGCTCCCTCA
1351 TCCCAGCAGC CTCGCACACT CCCTAGGCTT CTACCCTCCC TCCTGATGTC
1401 CCTGGAACAG GAACTCGCCT GACCCTGCTG CCACCTCCTG TGCACTTTGA
1451 GCAATGCCCC CTGGGATCAC CCCAGCCACA AGCCCTTCGA GGGCCCTATA
1501 CCATGGCCCA CCTTGGAGCA GAGAGCCAAG CATCTTCCCT GGGAAGTCTT
1551 TCTGGCCAAG TCTGGCCAGC CTGGCCCTGC AGGTCTCCCA TGAAGGCCAC
1601 CCCATGGTCT GATGGGCATG AAGCATCTCA GACTCCTTGG CAAAAAACGG
1651 AGTCCGCAGG CCGCAGGTGT TGTGAAGACC ACTCGTTCTG TGGTTGGGGT
1701 CCTGCAAGAA GGCCTCCTCA GCCCGGGGGC TATGGCCCTG ACCCCAGCTC
1751 TCCACTCTGC TGTTAGAGTG GCAGCTCCGA GCTGGTTGTG GCACAGTAGC
1801 TGGGGAGACC TCAGCAGGGC TGCTCAGTGC CTGCCTCTGA CAAAATTAAA
1851 GCATTGATGG CCTGTGAAAA AAAAAAAAA AAAAAAAAA AAAAAA
```
(SEQ ID NO:2)

FEATURES:

5'UTR: 1 - 131
Start Codon: 132
Stop Codon: 1332
3'UTR: 1335

Homologous proteins:
Top 10 BLAST Hits

	Score	E
CRA|108000024636236 /altid=gi|298316 /def=gb|AAB25526.1| transc...	793	0.0
CRA|108000024653390 /altid=gi|12742775 /def=ref|XP_009922.2| tr...	793	0.0
CRA|18000004926133 /altid=gi|339205 /def=gb|AAA61057.1| (L02648...	792	0.0
CRA|108000024042036 /altid=gi|12654675 /def=gb|AAH01176.1|AAH01...	792	0.0
CRA|18000004926130 /altid=gi|4507409 /def=ref|NP_000346.1| tran...	788	0.0
CRA|18000004926132 /altid=gi|339203 /def=gb|AAA61056.1| (L02647...	786	0.0
CRA|18000005170902 /altid=gi|7657639 /def=ref|NP_056564.1| tran...	561	e-159

FIGURE 1C

```
CRA|164000136745249 /altid=gi|11968124 /def=ref|NP_071979.1| tr...    554  e-156
CRA|18000005218941 /altid=gi|4572454 /def=gb|AAD23829.1|AF12128...    545  e-154
CRA|18000004926134 /altid=gi|4507407 /def=ref|NP_001053.1| tran...    128  1e-28
```

EST:
```
gi|10725490 /dataset=dbest /taxon=96...                               858  0.0
gi|5936410 /dataset=dbest /taxon=9606 ...                            835  0.0
gi|6888875 /dataset=dbest /taxon=9606...                             726  0.0
gi|6888872 /dataset=dbest /taxon=9606...                             726  0.0
gi|12258937 /dataset=dbest /taxon=960...                             686  0.0
gi|10947399 /dataset=dbest /taxon=96...                              680  0.0
gi|13287907 /dataset=dbest /taxon=96...                              680  0.0
gi|9121897 /dataset=dbest /taxon=9606...                             680  0.0
gi|13280819 /dataset=dbest /taxon=96...                              680  0.0
gi|8150776 /dataset=dbest /taxon=960...                              656  0.0
```

EXPRESSION INFORMATION FOR MODULATORY USE:
library source:
```
gi|10725490|  adult adrenal gland
gi|5936410|  adult uterus
gi|6888875|  adult head_neck
gi|6888872|  adult head_neck
gi|12258937|  adult lung_tumor
gi|10947399|  mammary gland
gi|13287907|  retinoblastoma
gi|9121897|  retinoblastoma
gi|13280819|  adenocarcinoma cell line
gi|8150776|
```

Tissue Expression:
Human hippocampus

FIGURE 1D

```
  1 MRHLGAFLFL LGVLGALTEM CEIPEMDSHL VEKLGQHLLP WMDRLSLEHL
 51 NPSIYVGLRL SSLQAGTKED LYLHSLKLGY QQCLLGSAFS EDDGDCQGKP
101 SMGQLALYLL ALRANCEFVR GHKGDRLVSQ LKWFLEDEKR AIDTAAMAGL
151 AFTCLKRSNF NPGRRQRITM AIRTVREEIL KAQTPEGHFG NVYSTPLALQ
201 FLMTSPMRGA ELGTACLKAR VALLASLQDG AFQNALMISQ LLPVLNHKTY
251 IDLIFPDCLA PRVMLEPAAE TIPQTQEIIS VTLQVLSLLP PYRQSISVLA
301 GSTVEDVLKK AHELGGFTYE TQASLSGPYL TSVMGKAAGE REFWQLLRDP
351 NTPLLQGIAD YRPKDGETIE LRLVSW
```
(SEQ ID NO:3)

FEATURES:
Functional domains and key regions:
PDOC00005 PS00005 PKC_PHOSPHO_SITE
Protein kinase C phosphorylation site
Number of matches: 2

1	75-77	SLK
2	174-176	TVR

PDOC00006 PS00006 CK2_PHOSPHO_SITE
Casein kinase II phosphorylation site
Number of matches: 6

1	67-70	TKED
2	90-93	SEDD
3	174-177	TVRE ·
4	226-229	SLQD
5	249-252	TYID
6	302-305	STVE

PDOC00008 PS00008 MYRISTYL
N-myristoylation site
Number of matches: 7

1	12-17	GVLGAL
2	57-62	GLRLSS
3	86-91	GSAFSE
4	149-154	GLAFTC
5	190-195	GNVYST
6	209-214	GAELGT
7	230-235	GAFQNA

PDOC00009 PS00009 AMIDATION
Amidation site

162-165	PGRR

SignalP results:

Measure	Position	Value	Cutoff	Conclusion
max. C	19	0.602	0.37	YES
max. Y	19	0.702	0.34	YES
max. S	5	0.974	0.88	YES
mean S	1-18	0.949	0.48	YES

Most likely cleavage site between pos. 18 and 19: ALT-EM

BLAST Alignment to Top Hit:
>CRA|108000024636236 /altid=gi|298316 /def=gb|AAB25526.1|
        transcobalamin II, TC II [human, endothelial cells,
        Peptide, 427 aa] /org=human /taxon=9606 /dataset=nraa
        /length=427
      Length = 427

 Score = 732 bits (1870), Expect = 0.0
 Identities = 376/427 (88%), Positives = 376/427 (88%), Gaps = 51/427 (11%)
 Frame = +1

FIGURE 2A

```
Query: 31    MRHLGAFLFLLGVLGALTEMCEIPEMDSHLVEKLGQHLLPWMDRLSLEHLNPSIYVGLRL 210
             MRHLGAFLFLLGVLGALTEMCEIPEMDSHLVEKLGQHLLPWMDRLSLEHLNPSIYVGLRL
Sbjct: 1     MRHLGAFLFLLGVLGALTEMCEIPEMDSHLVEKLGQHLLPWMDRLSLEHLNPSIYVGLRL 60


Query: 211   SSLQAGTKEDLYLHSLKLGYQQCLLGSAFSEDDGDCQGKPSMGQLALYLLALRANCEFVR 390
             SSLQAGTKEDLYLHSLKLGYQQCLLGSAFSEDDGDCQGKPSMGQLALYLLALRANCEFVR
Sbjct: 61    SSLQAGTKEDLYLHSLKLGYQQCLLGSAFSEDDGDCQGKPSMGQLALYLLALRANCEFVR 120


Query: 391   GHKGDRLVSQLKWFLEDEKRAI--------------------------------------- 456
             GHKGDRLVSQLKWFLEDEKRAI
Sbjct: 121   GHKGDRLVSQLKWFLEDEKRAIGHDHKGHPHTSYYQYGLGILALCLHQKRVHDSVVDKLL 180


Query: 457   -------------DTAAMAGLAFTCLKRSNFNPGRRQRITMAIRTVREEILKAQTPEGHF 597
                          DTAAMAGLAFTCLKRSNFNPGRRQRITMAIRTVREEILKAQTPEGHF
Sbjct: 181   YAVEPFHQGHHSVDTAAMAGLAFTCLKRSNFNPGRRQRITMAIRTVREEILKAQTPEGHF 240


Query: 598   GNVYSTPLALQFLMTSPMRGAELGTACLKARVALLASLQDGAFQNALMISQLLPVLNHKT 777
             GNVYSTPLALQFLMTSPMRGAELGTACLKARVALLASLQDGAFQNALMISQLLPVLNHKT
Sbjct: 241   GNVYSTPLALQFLMTSPMRGAELGTACLKARVALLASLQDGAFQNALMISQLLPVLNHKT 300


Query: 778   YIDLIFPDCLAPRVMLEPAAETIPQTQEIISVTLQVLSLLPPYRQSISVLAGSTVEDVLK 957
             YIDLIFPDCLAPRVMLEPAAETIPQTQEIISVTLQVLSLLPPYRQSISVLAGSTVEDVLK
Sbjct: 301   YIDLIFPDCLAPRVMLEPAAETIPQTQEIISVTLQVLSLLPPYRQSISVLAGSTVEDVLK 360


Query: 958   KAHELGGFTYETQASLSGPYLTSVMGKAAGEREFWQLLRDPNTPLLQGIADYRPKDGETI 1137
             KAHELGGFTYETQASLSGPYLTSVMGKAAGEREFWQLLRDPNTPLLQGIADYRPKDGETI
Sbjct: 361   KAHELGGFTYETQASLSGPYLTSVMGKAAGEREFWQLLRDPNTPLLQGIADYRPKDGETI 420


Query: 1138  ELRLVSW 1158
             ELRLVSW
Sbjct: 421   ELRLVSW 427
(SEQ ID NO:6)
```

HMM results:

Model	Description	Score	E-value	N
PF01122	Eukaryotic cobalamin-binding protein	829.9	8.6e-246	2
CE00052	CE00052 lymphocyte_transmembrane_protein_KAP	3.2	2.9	1

Parsed for domains:

Model	Domain	seq-f	seq-t		hmm-f	hmm-t		score	E-value
CE00052	1/1	1	11	[.	1	11	[.	3.2	2.9
PF01122	1/2	1	142	[.	1	143	[.	296.0	4.6e-85
PF01122	2/2	143	376	.]	197	450	.]	531.8	4.8e-156

FIGURE 2B

```
  1 MRHLGAFLFL LGVLGALTEM CEIPEMDSHL VEKLGQHLLP WMDRLSLEHL
 51 NPSIYVGLRL SSLQAGTKED LYLHSLMLGY QQCLLGSAFS EDDGDCQGKP
101 SMGQLALYLL ALRANWHDHK GHPHTSYYQY GLGILALCLH QKRVHDSVVD
151 KLLYAVEPFH QGHHSVDTAA MAGLAFTCLK RSNFNPGRRQ RITMAIRTVR
201 EEILKAQTPE GHFGNVYSTP LALQFLMTSP MRGAELGTAC LKARVALLAS
251 LQDGAFQNAL MISQLLPVLN HKTYIDLIFP DCLAPRVMLE PAAETIPQTQ
301 EIISVTLQVL SLLPPYRQSI SVLAGSTVED VLKKAHELGG FTYETQASLS
351 GPYLTSVMGK AAGEREFWQL LRDPNTPLLQ GIADYRPKDG ETIELRLVSW
```
(SEQ ID NO:4)

FEATURES:
Functional domains and key regions:
PDOC00005 PS00005 PKC_PHOSPHO_SITE
Protein kinase C phosphorylation site
                198-200   TVR

PDOC00006 PS00006 CK2_PHOSPHO_SITE
Casein kinase II phosphorylation site
Number of matches: 7
        1   67-70   TKED
        2   90-93   SEDD
        3   147-150   SVVD
        4   198-201   TVRE
        5   250-253   SLQD
        6   273-276   TYID
        7   326-329   STVE

PDOC00008 PS00008 MYRISTYL
N-myristoylation site
Number of matches: 7
        1   12-17   GVLGAL
        2   57-62   GLRLSS
        3   86-91   GSAFSE
        4   173-178   GLAFTC
        5   214-219   GNVYST
        6   233-238   GAELGT
        7   254-259   GAFQNA

PDOC00009 PS00009 AMIDATION
Amidation site
                186-189   PGRR

PDOC00428 PS00468 COBALAMIN_BINDING
Eukaryotic cobalamin-binding proteins signature
                165-178   SVDTAAMAGLAFTC

SignalP results:

Measure	Position	Value	Cutoff	Conclusion
max. C	19	0.602	0.37	YES
max. Y	19	0.702	0.34	YES
max. S	5	0.974	0.88	YES
mean S	1-18	0.949	0.48	YES

Most likely cleavage site between pos. 18 and 19: ALT-EM

BLAST Alignment to Top Hit:
>CRA|108000024636236 /altid=gi|298316 /def=gb|AAB25526.1|
        transcobalamin II, TC II [human, endothelial cells,
        Peptide, 427 aa] /org=human /taxon=9606 /dataset=nraa
        /length=427
        Length = 427

FIGURE 2C

Score = 793 bits (2026), Expect = 0.0
Identities = 399/427 (93%), Positives = 399/427 (93%), Gaps = 27/427 (6%)

```
Query: 1    MRHLGAFLFLLGVLGALTEMCEIPEMDSHLVEKLGQHLLPWMDRLSLEHLNPSIYVGLRL 60
            MRHLGAFLFLLGVLGALTEMCEIPEMDSHLVEKLGQHLLPWMDRLSLEHLNPSIYVGLRL
Sbjct: 1    MRHLGAFLFLLGVLGALTEMCEIPEMDSHLVEKLGQHLLPWMDRLSLEHLNPSIYVGLRL 60

Query: 61   SSLQAGTKEDLYLHSLMLGYQQCLLGSAFSEDDGDCQGKPSMGQLALYLLALRAN----- 115
            SSLQAGTKEDLYLHSL LGYQQCLLGSAFSEDDGDCQGKPSMGQLALYLLALRAN
Sbjct: 61   SSLQAGTKEDLYLHSLKLGYQQCLLGSAFSEDDGDCQGKPSMGQLALYLLALRANCEFVR 120

Query: 116  ------------W----------HDHKGHPHTSYYQYGLGILALCLHQKRVHDSVVDKLL 153
                        W          HDHKGHPHTSYYQYGLGILALCLHQKRVHDSVVDKLL
Sbjct: 121  GHKGDRLVSQLKWFLEDEKRAIGHDHKGHPHTSYYQYGLGILALCLHQKRVHDSVVDKLL 180

Query: 154  YAVEPFHQGHHSVDTAAMAGLAFTCLKRSNFNPGRRQRITMAIRTVREEILKAQTPEGHF 213
            YAVEPFHQGHHSVDTAAMAGLAFTCLKRSNFNPGRRQRITMAIRTVREEILKAQTPEGHF
Sbjct: 181  YAVEPFHQGHHSVDTAAMAGLAFTCLKRSNFNPGRRQRITMAIRTVREEILKAQTPEGHF 240

Query: 214  GNVYSTPLALQFLMTSPMRGAELGTACLKARVALLASLQDGAFQNALMISQLLPVLNHKT 273
            GNVYSTPLALQFLMTSPMRGAELGTACLKARVALLASLQDGAFQNALMISQLLPVLNHKT
Sbjct: 241  GNVYSTPLALQFLMTSPMRGAELGTACLKARVALLASLQDGAFQNALMISQLLPVLNHKT 300

Query: 274  YIDLIFPDCLAPRVMLEPAAETIPQTQEIISVTLQVLSLLPPYRQSISVLAGSTVEDVLK 333
            YIDLIFPDCLAPRVMLEPAAETIPQTQEIISVTLQVLSLLPPYRQSISVLAGSTVEDVLK
Sbjct: 301  YIDLIFPDCLAPRVMLEPAAETIPQTQEIISVTLQVLSLLPPYRQSISVLAGSTVEDVLK 360

Query: 334  KAHELGGFTYETQASLSGPYLTSVMGKAAGEREFWQLLRDPNTPLLQGIADYRPKDGETI 393
            KAHELGGFTYETQASLSGPYLTSVMGKAAGEREFWQLLRDPNTPLLQGIADYRPKDGETI
Sbjct: 361  KAHELGGFTYETQASLSGPYLTSVMGKAAGEREFWQLLRDPNTPLLQGIADYRPKDGETI 420

Query: 394  ELRLVSW 400
            ELRLVSW
Sbjct: 421  ELRLVSW 427
```
(SEQ ID NO:7)

HMM results:

Model	Description	Score	E-value	N
PF01122	Eukaryotic cobalamin-binding protein	906.3	8.6e-269	2
CE00052	CE00052 lymphocyte_transmembrane_protein_KAP	3.2	2.9	1

Parsed for domains:

Model	Domain	seq-f	seq-t		hmm-f	hmm-t		score	E-value
CE00052	1/1	1	11	[.	1	11	[.	3.2	2.9
PF01122	1/2	1	115	[.	1	115	[.	241.3	1.4e-68
PF01122	2/2	117	400	.]	145	450	.]	660.5	8.7e-195

FIGURE 2D

```
   1 ATATGTATGG GAAATATGCT GTCTTCCTAT TCCTACTCCC CCACCCTCTA
  51 GCACTGAGTC CAGGTAGGTA GGCAGGGGGG TGTCTCCCTC CTTTACTTCG
 101 ACACCCTAAC TACCTTGGGG ATCAGAAGTG ACTCTCTGGA AGGATGCTGC
 151 TGCTTCTCAC CAGAGGCTGA CGATAACGAA GGCTATCCTC CATGGCCACC
 201 TCCTCCAGGC TGCCTTCCTG GAAATAGGAA TCATAATAGT TGTTACTGGA
 251 AACAGGCAGA GGGTTGGGGG AGCCAAGGCA GTCCCACCCA GGACCAAGGT
 301 GGCTCCATTG CACACACTTC ACCATGACTC CCCTGAAGGT CCAAACGTGC
 351 GGTTCTGCGG AAGTTGGGCT CCCCACTGGC CTCCCTCAGT CCTCAGAACC
 401 TCCAGGGGTG CTCCTCCTAG TGGCCACATC CAGCCTTTCT GACTGGACAA
 451 CCTATCATTT AAAATTTTCA AGTAGTTCCG TAAACAGACA CACGTTGCTG
 501 TATTTATTTA TGTCAAGGGC TTGGTTTGTG ATAAGTCAGG CTCAAAAAGA
 551 TTGTCTTAAA AGAGTGAACC TTGGCAATTT ACCATAAAAT AATTGCAATG
 601 CAGATTGTGC ATGGAAATGA TTGGAGATAT TTTAAGGTCA TAGTGTCTTC
 651 ACAAATTGAG CTGAAAGGGA ACTGTTAGGA TGATCTTGCC TAACCCTCTC
 701 ATCTCACACA GGAAGAACTA TTTTAAACTC GAGAGGTTAA GTGACCTGGC
 751 CAAAGTCACA CAGCCACCAC TAGTTAACTC GTATACATTG ATTCTCCTGT
 801 GGGGCTGGGC AGATGAGGAA TCTTTTGTTC TCTTCCCTGT TTGCAGAGAT
 851 TTTTTTTGAG GTTACTTTCC GAGTTCTGGC AAGTACCCCT GCTTCTGGTA
 901 GCTTTGTGTC TCGATTCAAT CTCATTCTTT TTATTTTATT TTATTTTTGA
 951 GACAGGGTCT CACTTTGTCA CCCAAGCTGG AGTGCAGTGG TGTAATCTTG
1001 GCTCACTGTA GCCTCCACCT CTTGGGTTCA AGCGATCCTC CTGCCTCAGC
1051 CCCCCAAGTA GCTGGGATTA CAGACGTCTG CCACCCAGCC AGGCTAATTT
1101 ATGGTTTTTT GTATGTGTTT TTTGTGTTTT TGTAGAGACA GTGTTTCCCC
1151 ATGTTGCCCA GGCTGGTCTC CAACTCCTGA GCTCAAGTGA TCTGCCCGCC
1201 TCAGCCTTTC AAAGTGCTAG GATTACAGGT GTGAGCCACC GTGCCCGGAC
1251 TTAATCCCAT TCTTTAACTT GTTTTGTTTT GTCCTCTCCA GGAGGCTCCC
1301 AGCCCTTTCG GATTGGTTGA GAAAAGTGGC CTGGCTGGTC TGGGGCCAGC
1351 AGCACCCACC CTCCCCTCAA TTGCCCAACT CCCCCCCCCA CCGAACTGCC
1401 CAACTCCCCC TCCCCAACTG CCCAACTCCC CCACCCCCAC AATCCCCTCC
1451 CGCCACAACT GAGGGAGGCG GTGCTGAAAA ACAGCTCAGT CCAGCAATGC
1501 TGCTCACGTG ACCACTGCAG CTGCAGCTCC CGTTCCACTC CTTGTCCTGG
1551 GCTAGGTGGG CACTACCAGG GGCTCCTTTG GTAAGGAGTA CCGGGTAGGC
1601 ACCCGGTCCT GCCAATCCAC CACTGGAACA GCTGGGGGGA CAGCAGACAG
1651 GCACGGTCGG ACAGACTTGA CAGATCAGGC ATCAGGCCCT CTGCGCTGGT
1701 CCCGGGCTCT TTAAGCAGGA ACGTGAATGG CCTCAAGATG TCTCACATGG
1751 TCCCACTAGC CCTCCTCCTC CCTTTGTTCC CTACCTCCAG GAGGGCTGCT
1801 CTGCCCTTCC TTCCTCTGTT CTTTGGCCTT ATGTTCCCCG CCACCACAGG
1851 CCTTCCCCCG CCCCACCCCT CTGCAGACTT AGCCGTGCAT TGCAGGCATG
1901 GAGGATTAAT CAGTGACAGG AAGCTGCGTC TCTCGGAGCG GTGACCAGCT
1951 GTGGTCAGGA GAGCCTCAGC AGGGCCAGCC CCAGGAGTCT TTCCCGATTC
2001 TTGCTCACTG CTCACCCACC TGCTGCTGCC ATGAGGCACC TTGGGGCCTT
2051 CCTCTTCCTT CTGGGGGTCC TGGGGGCCCT CACTGAGATG TGTGGTGAGT
2101 AACTCGCCTC TATCCTGTGC CTCTTTCCTC CTGGGTCCTT AGTGGGGTGG
2151 CTAGGGCATA GGATGAGGGA ACTTACCTGC CCTTCTAAGC TCCCATAGCA
2201 GTTTGGGCTT AGCTGGACCT CAGCATTTAA CACATCCTAT TGTGATTGAT
2251 TATATGTTTG ACTCCTCACC AGACAAGATC TCCGTTAATT CAGTCATTCG
2301 TTCACACATT CATTCAGCGC ATACTGAGCC TTTTCTGTGT CAGGCCCAGT
2351 GTTAGCCTTT GGGGAACGTG CAAAGCATGA GACAAGTCTA ATCCCTGCCA
2401 TCCTAGAGCT TATGTTCTAG GGAAGGGGGA CAGACAAAAG AAATGGTTAG
2451 GTGCTCCCAC CTGAAATCTC AGCATTTTGG AAGGCTGAGG CGGGAGGGGA
2501 GGATCGCTTG AGCTCAACAG TTCAAGGTCA GCCTGGGCAA CATAGGGAGA
2551 CCCCATCTCT ACAAAAAATA AAAAAAATTA AAAAATAGCT GGGCATGGGG
2601 AAGACTTTCT GAAGACCAAG AGGACACATG GGAGCTGAAA CTCGAAGGAA
2651 GAAAAGGAGC TGGCAGGAAA GGAGTGGGGG ACACACATTC TAGGCAGCAG
2701 GAAGTGAGCC TTCGGAGGTC CTGCCTGCTC CAGCTCTGTG CCCCAAGGGG
2751 TCTCTTGGAG CACAGTCTCC TGGGACCTGT CTATGAGTCT GAGCTTAGAG
2801 GCTCAGGGCT GCTCCTTCAG ACAGGAGGCA GAAGGCAGAC TTTGGGAACT
2851 TTGGGCCGCC CACGCGCCTT TTCTCCTCCT CTGCACCTAG GATTACGTTG
2901 AGCAATACAC TTTCACCCCC ATGGTCTCTT GAGACCCTGG GGAAACCCTG
2951 AGAGGTGGGT GCAGTCATGT CCAGGTGTCA AGTGAAGAAG TCGAGGGTTG
3001 GAGGGGCTGA GTGACCCACT CAGGGTGCTC CACCTTTTCC AGAGCTTTGC
3051 TGAACTTAGT TTTTAGAACT TGAAGCCTCG TTTGTTTTCG TTTTGTTTTT
3101 TGTTGAGAGA GGGTTCTCCCT CTGTTGCCCA GGCTGGAGTG CAGTGGCACG
3151 ATCTTGGCTC ACTGCAGCCT CTGCCTTGTG GGTTCAAGTG ATTCCCCCAC
3201 CTCAGCCTCC CAAGTAGCTG GAGACTGCAT GTGCATACTA CCATGCTTGG
3251 CTAATTTTTG TATTTTTTTG TAGAGACAGG GTTTCGCCAT GTTGCCCAGG
3301 CTGGTCTCGA ACTCCTGGGC TCAAGTGAAA CTCTTGCCTC GGCCTCCCAA
3351 ATTGCTGAGA TTACAGGCGT GAGCCACCGT GCCCGGCCAG AACTCCAAGC
3401 CTCTCATCTG TGTTCCATAA ATGCAATCAG ACACCTCAGG TCTGGGCCCA
```

FIGURE 3A

```
3451 GGAACCCCAG CTCTTGGTTC ATGTCCGGAC AGTCCCCAGG GGAGTTCTGG
3501 GTTCAACCAG CAAGAGCTCT TCCTCCTGGC TGATCTGGTC CTCAGCCTTG
3551 GACAGTTAGT CCATTAACCT GACCCCACAG GAGCCCCAAT CCCTTGGGGT
3601 CTGGGGAATC TTGAACTGGG GTTTGGGGTG CAAATATCTG CACTGAGTCA
3651 CTTAATTGCA CCCAGCCTCA TTCCTTTATC TGTAAAGTGG GCTAAGAATG
3701 CTCCCCTGCC TTCCTCCTCG GTGTAGTACG AGGAAGGATC CCATGACACC
3751 TGCTCTCCCA GTTTAAAGCT CTATATGTAT GTTGTGAAAT TGACAGGGAT
3801 CGCTGCACAA ACGCTAATGC AAAGTGGGCT CCTGTGCTTC CTTTTCTCTT
3851 TCTTCTTCTT TTTTTTTTTT TTAATTTTCT TCTAGAGATG AGGTCTCACT
3901 ATATTGCCCA GGGTTGGTTT CAAACTCCTA GGGTCAAGCG ATCCTCCCAC
3951 CTTGGCCTCC CAAACTGCTG GTATTACAGG CGTGAGCCAC TCTGTCTGGC
4001 TCCTATGCTT GTGAATGTCA ACAGCAATCA GCCCTTAGCT GGCAGGGCTG
4051 GGTTGGTAGG GCGAGAGCTC ACCCAAGGCT GCTTTTATTA CCCTGCGTGA
4101 ATCTGCCTGG CCCCTTCCTT CTAAGGAGGT TGCTCTGTGG TTGTCAGTCT
4151 CTCCCTTTAC AGCTGGATCC TGATCTTTCA GTTTCTAACC CTGTGCTGAC
4201 TCATCGTGCT GGAAGTGAGA GCCCGGGGTG AGGTCAGGGA ACTCCCTTGC
4251 GCGTTTCAAG AAAAGGGAAA AGGAAAGAGA GGTGAGGAGG GGGCAGATG
4301 ACCAGAGAGA CACAGGCTGA GAGAGACTGA GACAGACCCA GAGAGCCTCA
4351 CACATTGAGT GACAGAGACG GAGAAATGGA GATAGGCACC AAAAAATGGT
4401 TCTCAGTGAC AGAAAGGGAA AAAAGCAACC CCCCAGTCTC TCTTAACATC
4451 TGGTGAGAAA CCAGCCATGT GCTTTGGTCT GGGCCCACAC AGCAAAGGAT
4501 TATGTAGGGT TTCATGCTGG TGGATGGTCA CCTTATAGCA ACAGGTATCT
4551 GGGGCTGTCG GGAAAACAGA CACGAGGTTG TGGGACCCAG ACCCACAGAG
4601 ATGGAGCTGT TCTAGGAGCT CTGGTCCTCG TTCTGGTCCC CTGGGATATG
4651 GCACAGTGAA GGCCACCATC AGGCAGCTGG AGCCCAGCAG CAACTGGGAG
4701 GCAGTAAACA GGGACCGAAA GTGCAAGGTT ACCTCCGAGG CAAACTACTC
4751 TAAGCTACCC TGTGCTGAGC TCAAGTCCCT TGGAACTATC CCTAAGGCTT
4801 CCGCTTCCAG AGTGTTTGAG TATTTTCGTT GCACAGCTTC GAATAAATCC
4851 CACAGCAACA GGTAAACGGC TGCAAGCTGT GACTGTTTTC TAAGAGCTCA
4901 TCTCACAATC TCAGGTCCTC TTCATTTAAA CAGAGATGGC AGGAAAGGCG
4951 TTATTTTGAG ATCTGCATGG AGGAAGTTCA CCAGGCAGCC TCAATTCACC
5001 AGCTGGAAGT TTGCGTTGTT TGGAAATTTG ATGTGTAACA CGTTCTGCAT
5051 GTGGGCTGAT GTTTTTGTAA ACGGGTAGCA CACACATTCA GCAGGGCACC
5101 AAAGAGCGGG GGCTTTGCAG TTAGGTCCAT CCTTGGCTCT GCAGCCTTGT
5151 GTAAGACATG ACACGACTTT GAACTTCTGT TTCCTCTTCT GTGCAAAGCA
5201 ATGATGACAG TATCTACATC ACAGGACTGG CATGAGGACC AAGTGAGATT
5251 GGGCAAGGTG CCCGGGCACA CCAGTCTCAC TGTCACTGCT GATGGGCAGA
5301 GTGGTTGCCT GGCAGTAGCA TCCTCTATCT TCAGCCCACC ACCTCTCTTG
5351 CTGGCTCACT CCAACTGCTC TTTAGAGATA CACGCTTCCC CTCTTTTCTC
5401 CTCCCACTGC CTTTCAGTAT GGCTGCATTT CCCCCTGCAA GTTGGTGTGT
5451 GCTGGGTGGA GGTGGGGGTG AGGACATGTA TTCTCTGGAG AAGGCCCTGG
5501 TAACGTCAAA GCACTTCTTT GCTGGTGGCC TGGCCCTGTG ACCTCATTTG
5551 TACCATTTTC TTTTCTAAGA AATACCAGAG ATGGACAGCC ATCTGGTAGA
5601 GAAGTTGGGC CAGCACCTCT TACCTTGGAT GAACTGGCTT TCCCTGGAGC
5651 ACTTGAACCC CAGCATCTAT GTGGGCCTAC GCCTCTCCAG TCTGCAGGCT
5701 GGGACCAAGG AAGACCTCTA CCTGCACAGC CTCAAGCTTG GTTACCAGCA
5751 GTGCCTCCTA GGGTATTGCC ACACTCTCTT TTTCCATGTC TTGCTCCACA
5801 TACTAAGAGA TGGGAAACTT GGGTACTAGT TTGGGCCTGT CACCACTTTG
5851 TGGGCAGACC TTAGGCAAAT TTTCTCCATC TATAGAATGG AGGACCTTTG
5901 TCCATCTATA GAATGAAGGG GTTGGTTGGA TTAGATCAGA GATGCTAATG
5951 CAAGGCTCCT TTTGCTACTA CTGTCCATCA TGTGTCTGAG GCAGACATAA
6001 CTAATCCGTG ACTATACTCT TTGATGATGA GCCCAGGAGC AGCATCTGAC
6051 TCTATGCTCC CTTAGTGTGC CTGAGGCAGA TATCACTAAT CGATGACTGC
6101 AGTCTTCTAC ATTGAGCTTA GAAGCAGCAT CTGACTCTGT ATGCTCCCCT
6151 CCCATGCATG AGGCAGACAT CAGTAATCCA TGACCGCATT CTTTCATACT
6201 GAGCCCAGAA GCAGCATCTT TTCTTTTCTT TCCTCTCACT CTGTTGCCCA
6251 GGCTAGAGTG CAGTGGCACA ATCTTGGCTT GCCCCAACCT CCAATTCCCG
6301 GGTTCAAGTG ATTCTCGTGC CTCAGCCACC TGAATAGCTG GGATTACAGG
6351 CGTGTGCCAC CATGCCCAGC TGATTTTTGT ATTTTTGGTA GAGATAGGGT
6401 TTCACCATGT TGGCCAGGCT GGTCTTGAAC TCCTGACCTC AGGTGATCCG
6451 CCTGTCTTGG CTTCCCAAAG TGTTGGGATT ATAGGCATGA GCCACTGCAC
6501 CAATCCAAAA GCAGCATCTT TGTGCTCCCT TTTCAAGAGG CATCACAGAG
6551 AGGCCTGTTT TGGGGTTTGA ATGAGAGGCG AAGAATCAGC CATGGAGTGC
6601 CTCTTTCTCA GACTCCCTCT TGAGAAGTGG GTGCAGGGGT GGAGAGAAAA
6651 GAAGACTAGG CATAGTGGCT CATACCTGTA ATCCCAACAT TTTGGGAGGC
6701 TGAGGCAGGA AGATTGCTTG AGCTCAGGAG TTTGAGACCA GCCTAGGCAA
6751 CATAGTGAGA CCACATCTCT TAAAAAAAAG AAAAAGAAAA AAAATGAGCC
6801 AGGTGTAGTG ACTCATGCCT GTGGTCCCCA CTTCTCCGGA GGCAAGGGTG
6851 GGAGGATCTT TTGAGGCTGA GAAATCGAGG CTACAGTGAG CCATGGTGGC
```

FIGURE 3B

```
 6901 ACCACTGCAC TCCAGCCTGG GAGACAGAGA GACCCTATCT CAGTAAAAAA
 6951 AAAAAATAAA AATATGGCTG GGTGTGGTGG CTCACGCCTG TAATCCCAGC
 7001 ACTTTGGGAG GCCAAGGTAG GTAGATCACA TGAGGTTAGG AGTTCGAAAC
 7051 CAGTCTGGCC AACATAGTGA AACCCTGTCT CTACTGAAAA TACAAAAAAT
 7101 TAGCCAAGGG TGGTGGTGGG CAACTGTAAT CCCAGCTACT TGGGAGGCCG
 7151 AGGCAGAAGA ATCGCTTGAA CTCGGGAGGC GGAGGTTGCA GTGAGCTGAG
 7201 AACATGCCAC TGCACTCCAG CCTGGGCAAC AAGAGCGAAA CTCTGTCTCA
 7251 AAGAAAATAA ATAAATAAAA TAAAAAAATA AAAAAGGAGG GGGCATATGG
 7301 GTGAAGTATG GACAAAATAG TGGGGCAGGC ACAGATGATC TGGACACAGG
 7351 AGCCCTTGGA GTTTATTCTT GAATCTAACT GTTCATCTTT ATTAAATATT
 7401 TGTGGCATAC ACCTCACAAC AACATAGCCA ACACACCTCC TTTTGGAGCT
 7451 TTTATCGAAG TTTCCCACTG TTAAGATTTT TTCCCGCTTT GTGATGCGGG
 7501 TGGGGTGGGT GCTGTAAGCA GGCTTACGGG GTGGCAGTTT CTCACAAAGG
 7551 CATTAACTGG CCTTGTCCTA GGTCTGCCTT CAGCGAGGAT GACGGTGACT
 7601 GCCAGGGCAA GCCTTCCATG GGCCAGCTGG CCCTCTACCT GCTCGCTCTC
 7651 AGAGCCAACT GTGAGTTTGT CAGGGGCCAC AAGGGGGACA GGCTGGTCTC
 7701 ACAGCTCAAA TGGTTCCTGG AGGATGAGAA GAGAGCCATT GGTGAGCAGA
 7751 CACCATCCGC TGGGGGTGGG GAGCAGCTGG GAGGGCTCAT CAGATGATAT
 7801 TCTCCAATGA GAATCAGAAC TTTGGGTTTT CTCCCCAGGC GTCTTTCCCA
 7851 CCATCCATTC TGCCCATCTC ACTGCCTACG TAGAGGCTCG AACCTGTCCC
 7901 CATAGCCATC CTTGACCCAG CTTTTCCCGC GCTGCACACA TACTATTGAC
 7951 AGGTGTGTTT CGTGGTTTTT TGTTTTTTGT TTGTTTGTTT GTTTTGAGTT
 8001 GGAGGTTTGC TCTTGCTGCC CAGGCTGGAG TACAATGGCG CAATCTCAGC
 8051 TCACCGCAAT CTCTGCCTCC TGGGTTCAAG CAATTCTCTT GCCTCAGCCT
 8101 CCTGAGTAGC TGGGATTACA GGCATGCGCC ACCACACCCA GCTAATTTTG
 8151 TATTTTTAGT AGACGTGGGG TTTCTCCATG TTGGTCAGGC TGGTCTCGAA
 8201 CTCCTGACCT CAGGTGATCC GCTTGCCTTA GCCTCCGAAA GTGCTGGGAT
 8251 TACAGGCATG AGCCACTGCG TTAGGCCCAC TGACAAGCCT TGTATTGGCT
 8301 AGCCACCAAG ATTGACTTGA TTATCCACCT TCGGGACAAC TGGACAGCCT
 8351 GCTTATGACT TACGCCATAG TCTGTCTCTA CTAGCTCTCC TGCCCTGACT
 8401 TGACCCAGCA TACAACAGCC AGAGCCAGCC TTTTCAATAT AAACCTGATC
 8451 TTGCTGGCAC TGCTTAAACC CTGCAGGGGC CTCGCACTGC TCCATGGCCC
 8501 AGCCTGTCTA CCCTTACCTT CTGCCCAGGC TCTGCTCATC CATTCTCTGC
 8551 CTCCCACACA CCTGCCCTCT GTGGGCTCCA GCCATACCAT CTCTCAACTC
 8601 ATAAGCCAGT TTTTTCATAC AGGCTCCCTC CATCTGGACT GGCTTCCCTG
 8651 CGTGCAGTTC ACTCCTGCTC TACCTTTGGC TCTGCCTCCA CCCATCCTCA
 8701 GCCGTCTCCA GCATTACCTC CTTGGAGAAT CCTGCCTTGA CTTCCCAGCC
 8751 ACCCAAATAT CACTACTTGG TCTGCATTCT CGTTGCAATT GCAGTCGCAT
 8801 GAGCAATTGC TGTGGTTGAG GCCCGAACTG CGCAAGTGCC TGTCTGCCAT
 8851 GGGTCTCCTG CTTCCTCTAA GCACAGTGCC TGACACACAG TGAGACCTCA
 8901 GCACGTATGG GCTGAGGCAA TGAAGGAATG AAGGATCCCA TGACCCAAAA
 8951 GAGCCTGTTG GAAAGTGCAG GCCAGGGTCC CAGGTGCTGG CGGGGCTGGC
 9001 TGCTGGGTGG GGGCAGAGAG GCAACCCCTC TGTTTTTTTC CCTCTCAGGG
 9051 CATGATCACA AGGGCCACCC CCACACTAGC TACTACCAGT ATGGCCTGGG
 9101 CATTCTGGCC CTGTGTCTCC ACCAGAAGCG GGTCCATGAC AGCGTGGTGG
 9151 ACAAACTTCT GTATGCTGTG GAACCTTTCC ACCAGGGCCA CCATTCTGTG
 9201 GGTGAGTAGG TCAGACCGTG CCAAGGCCAG GCTGGCACTC CCTCAGTCCC
 9251 CAGGTCTGCA CTGATGACGT CCATACCCTG GCCCCACAC TCACCTTTCC
 9301 TTGGGGCTCC TCCGAATCAA GTCCTTTAGG GACGAATTGG CGAGGGCTCA
 9351 TGGGTGATGC TCCAGCTGTG AGCCAGCTTT GGAGCTGGTA GGTGGATCTC
 9401 TTGAGGCCAG GAGTTCAAGA CAACGTGGTG AAACCCCATC TCTACTAAAA
 9451 ATAAAAAAGT TAGCCGGGCA TGGTGGCACA TGCCTGTAGT CCCAGCTACT
 9501 CGGGAGGCTG AGGCAGGAGA ATCACTTGAA CCTGGGAGGC GGAGGCTGCA
 9551 GTGAGTGGAG ATCGCACCAC TGCCCTCCAG CCTGGGCAAC AGAGTGAGTG
 9601 AGACTCTGTC TCAAAAAATA AAAATAAAA TAAAACTCCC CTAGTGATTC
 9651 CAATGTGCAG CTAAGTTTGG AAATAGGTGG TATGGGGTCA AGTCCTCTTG
 9701 GGCCTCCCTC CTCCAGTCCT TCTCCCTAAC CTCTAGCCCT CAAGTTGCAG
 9751 AGTGATCAGC CAAACCAGTT TGCCCAGAAA TGAGCAGTTT CCTGGGACAC
 9801 AGGATTTTCA GAGTCCAGAC AAGGAAAGTC TTGGGCAGAC CAGGTTGAGT
 9851 TGGTGCCCTT AGCTGATCTG ACCATGTTGC CCTTCTTCTC CAAGCCCTCC
 9901 TGTGGTTGTC CATAGCTACA AGGGCCTGAC CCTCAAGCCC CTGCCTGTCC
 9951 TGGCCCCTTT GGCTCTCCAG CTCATTGCAT GTTCTGTCCC CCACTTCAAG
10001 ACACAGCAGC CATGGCAGGC TTGGCATTCA CCTGTCTGAA GCGCTCAAAC
10051 TTCAACCCTG GTCGGAGACA ACGGATCACC ATGGCCATCA GAACAGTGCG
10101 AGAGGAGATC TTGAAGGCCC AGACCCCCGA GGGCCACTTT GGGAATGTCT
10151 ACAGCACCCC ATTGGCATTA CAGGTGGGAA AGAGACCCTG GAGCCATGGC
10201 CACCCTGGGG AACAGTCGGG TGGAGTGGTC AGGTGCTGGA ACACCTAGCC
10251 CCTCCCTGCC GGCTGACCTC CTCTCTCTCT TCCTCACTCT ATCACCAGTT
10301 CCTCATGACT TCCCCCATGC CTGGGGCAGA ACTGGGAACA GCATGTCTCA
```

FIGURE 3C

```
10351 AGGCGAGGGT TGCTTTGCTG GCCAGTCTGC AGGATGGAGC CTTCCAGAAT
10401 GCTCTCATGA TTTCCCAGCT GCTGCCCGTT CTGAACCACA AGACCTACAT
10451 TGATCTGATC TTCCCAGACT GTCTGGCACC ACGAGGTAGC CCAACTTTTT
10501 GTGGAAGCAC AGCCCTTTAC AATCTGCTGC GCACCCATTG ACGTCCCAGT
10551 GAGGGGAGGT TGCTTCATCC TGATTTGCTG AGTCAGCACA AGTTTGTGGG
10601 TGTGCATGGG ACACAGTAGC CAAAATGTGG TCATAGCTTC TAGAAGCTCA
10651 CAGTGTGGGG AGGAAGACAG TAAATGGAGA TCCCTGGGCA TATCGCTTGT
10701 GTGATACCCA GTACAGAAAT GTTTGGATGG ATGGATGGAT GGATGGATGG
10751 ATGGATGGAT GGATGGATGG ATGAGGAGAG ACACATTTTG GTTAACTCTA
10801 ATACAACATG ATAAGCCCCA GTAGCAGCAT GATCCAGGCT TTCTCTGAGA
10851 GAGGGTCTGA GGACGTGACT GGGATTTGCC AATTAAGAAT GGAGAAAGAG
10901 GCCAGGTGCA GTGACTCATG CCTGTAATCC CAACACTTTG GGAGGCCGAG
10951 GCGGGTGGCT CACCTGAGGT CAGGAGTTCG AGACCAGCCT GGCTAACATG
11001 GCGAAACTCC ATCTATTAAA AATACAAAAA AGTAGCTGGG TGTGGTGGCG
11051 AGTGCCTGTA ACCCCAGCTA AGCTACTCAG GAGGCTGAGG CAAGAGAATC
11101 ACTTGAACCT CAGAGGTGGA GGTTGCAGTG AGCCAAGATC ATGCCACTGC
11151 ACTCCAGTCT GGGTGACAGA GTAAGACTAT GTCTCAAAAA AAAAAAAAAA
11201 AAATGGAGAA GAAGGAAGCT GGACATGGTG GCTCGTGCTT ATAATCCTAG
11251 CACTCTGGGA AGCTGAGGCA GATGGATTGC CTGAGCCCAG GAGTTTGAGA
11301 CCAGCCTGGG CAACATGGTG AAACCCTGTC TTTACTAAAA TACGAAAGAT
11351 TAGCCAGGCA TGGTGGTAGA CACCTATAAT CCCAGCTACT AGGGAGGCTG
11401 AGCCACAAGA ATCACTTGAA CCTGGGAGAC AGAGGTTGCA GTGAGCCGAG
11451 ATCGCGCCAT TGCACTCCAG CCTGGGCGAC AGTGTGAGAC TCTGTCTCCA
11501 GAAAAAACAA GAATGGATAG AGTGGAGCCA AGAAGAGGCA GGAAGAACAA
11551 AGACACAGAG GTGCACAGAG TTTGGGGGAA TTTTGAGGAA TGGTCTTGCA
11601 AAAGAGTGGG ATCTGGGAGA ATGAGTGGGA GTGGAAAGCA GATGAATGAA
11651 GAGAAGGTGA GCGCATCAGG GTAACAGAGA TGCGTTGTGA ACAAATGCAT
11701 GTTCTAGGAA GAGCCCTCTG GAGTGCTAGG TGCCAGAGAG GTGGGAGGAA
11751 GGATACTGGA AGCAGAGAAA CCAGTGAGGG GCCTGATCTT GGGTGGTGGG
11801 GAATGAGGGA CAGGGGAGGC CGGGATGGAA GCCAGGTGGT GGGGAATGAG
11851 GGACAGGGGA GGCCGGGATG GAAGCCAGGT TTCAGCTGAG CAGGTGGCGG
11901 TGGCATTGAT GGAGATGAGG ACATGGGGAA GGACAAAGTC CAGGTGTCCT
11951 TGAGGGAAGA CAAGAAGACA AATAATCCAG GCTCTCTGTC CTCACACCAG
12001 CTGCCCGCCC CTTTCTTCCT GGCACAGTCA TGTTGGAACC AGCTGCTGAG
12051 ACCATTCCTC AGACCCAAGA GATCATCAGT GTCACGCTGC AGGTGCTTAG
12101 TCTCTTGCCG CCGTACAGAC AGTCCATCTC TGTTCTGGCC GGGTCCACCG
12151 TGGAAGATGT CCTGAAGAAG GCCCATGAGT TAGGAGGATT CACGTGAGAC
12201 TCCCACCTCC CAGTCCTCAC CCCACCCAAC CTCACATGCC TGATAACAGG
12251 GTCACAGAAA AGACGGGGAA CAGAGGAGAG GGTTCCCTCG GGAGAGACAC
12301 TGGCCCTGCT TCTGCTTCTA CCTGCTCAGC TCCTTTCTTG CCCACGGTGT
12351 TATGGAAACA GGGAGCCATA GGCCAGCATT GTCACTGAGA GAGCAGGCTT
12401 TGGAGGCAGA GCCCCCCAGT TGGAATCCCA ACTCTAACCA GCTAGGTTCC
12451 AGGTAGGCAC CCACAATTCA CCGAGGAGAA CAGTTGTGCC CCTTCCCTGC
12501 AGGGCCAGTG TGAAGAGTCC AGGAGTTAGT ACACATAGAG ATAGTGGCAT
12551 GTGCTTTTTA TATGTGCAAG GTCCAGCACA TAGCAAGCGC TCAACACAGC
12601 GTTGCTTTCA TCAGAGTAAG AACTGTTTTT TGTTTGTTTG TTTGTTTGTT
12651 TTTAAGAGAC AGGGTCTCAA TCTTATCACC CAGGCTGGAG TGTAATTGTG
12701 CAATCACGTC TCACTGCAGT CTCGAACTCT GGGGATGAAG CAACCCTACT
12751 GTCCTGCCTC AGCCTCCCAA ATAGCTGAGA CTATAGGCAC GTGCCACACA
12801 ACCCTGGGTA ATTTTTTTTT TTTTTTTTTT GAGATAGGGT CTCTGTCTGT
12851 TGCCCAGGCT GGTCTCAAAT TCCTGGCCTC AAACCATCCT CACACCTGAG
12901 GCGCTCAAAA TATTGGGATT ATAGGTGCGA GCCATCATGC TCAGCCAGAA
12951 TAATAACTGG TTTTTTTTGT TTTTTTTTTG AGACAGAGTC TCACTCTATT
13001 ACCCAGGCTC TGGAGGCCCA ACTCGTGTTT GTGTATTTGT TTATTTTTAT
13051 TTATTTATTT ATTTCGAGAC AGAGCCTCTC TCTTTCACCT AGGCTGGAGT
13101 GCAGTGGCGC AATCTCGGCT CACTGCAACC TCCGTCTCCT GGGTTCAAGT
13151 GATTGTCCTG CCTCAGCCTC CTGAGTAGCT GGTGCTACAG GCGCGTGCCA
13201 CCATGCCCAG CTAATTTTTG TATTTTTAGT AGAGACAGGG TTTTACTATG
13251 TTGGCCAGCT GGTTTCTAAC TCCTGAACTC GGGTGATCTG CCTGCCTCGG
13301 CCTCCCAAAG TGCTGGGATT ACAGGCATGG GCCTCCGTGC CCGGCCATGT
13351 ATTTATTTAG GCAAGGTCTC TCTCTGTTAT CCAGGCTGAA GTGCAGTGGC
13401 ACATTCATAG CTCACTGCAG CCTCAAATTA TCCAAGTAAC AGGGACTACA
13451 GGCATGCACC ACCACACCCA TCTACTTTTT TTTGAGATGG AGTCTCCCTC
13501 TGTCGCCCAG ACTGGGTTGC AGTGGCACAA TTTCAGCTCA TGGCAGCATC
13551 TACCTCCCAG GTTCAAGCGA TTCTCCTTCC TCAGTCTCCC GAGTAGCTGG
13601 GACTATGGGC ATGCACCACC ATACCTGGCT AATGTTTATA TTTTGAGTAG
13651 AGATGGAATT TTGCCATTTT GGCCAGGCTG GTCTTGAGCT CTTGACCTCA
13701 AGTGATATGT CTGCCTCAGN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
```

FIGURE 3D

```
13801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
13951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
14951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
15001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
15051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
15101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
15151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
15201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
15251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
15301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
15351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
15401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
15451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
15501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
15551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
15601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
15651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
15701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
15751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
15801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
15851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
15901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
15951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
16001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
16051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
16101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
16151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
16201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
16251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
16301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
16351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
16401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
16451 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
16501 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
16551 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
16601 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
16651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
16701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
16751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
16801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
16851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
16901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
16951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
17001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
17051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
17101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
17151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
17201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
```

FIGURE 3E

```
17251  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
17301  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
17351  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
17401  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
17451  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
17501  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
17551  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
17601  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
17651  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
17701  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
17751  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
17801  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
17851  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
17901  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
17951  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
18001  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
18051  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
18101  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
18151  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
18201  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
18251  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
18301  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
18351  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
18401  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
18451  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
18501  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
18551  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
18601  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
18651  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
18701  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
18751  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
18801  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
18851  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
18901  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
18951  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
19001  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
19051  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
19101  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
19151  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
19201  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
19251  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
19301  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
19351  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
19401  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
19451  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
19501  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
19551  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
19601  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
19651  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
19701  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
19751  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
19801  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
19851  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
19901  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
19951  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
20001  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
20051  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
20101  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN
20151  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNNNNNNNN  NNNCCAAATC
20201  AACCAGTTGC  ATAAATCACT  CCTCTATCTT  CCTTGGGGTG  GAAAGTGGAT
20251  GGGAGTTATA  ATTTGAGTTC  TCTTTTGTCT  TAGTCCATTG  AAGCTGCTAT
20301  TACAAAATAC  CATAAACTGG  GTGGCTTATA  AACAGCAGAA  ATGAGGCCGG
20351  GTGCGGTGGC  TCATGCCTAT  AATTCCAGCA  CTTTGGGAGG  CCAAGGCAGG
20401  TGGATCACCT  GAGATCAGTA  GTTCAAGACT  AGCCTGACCA  ACATGGTGAA
20451  ACCCTGTCTC  TACTAAAAAT  ACAAAAAATT  AGCTGGGGGT  GGTGGCGGGC
20501  ACCTGTAATC  CCAGCTACTC  AGGAGGCTGA  GGCAGGAGAA  TCGCTTGAAC
20551  CCAGGAGGCG  GAGGTTGCCG  TGAGCTGAGA  TCACGCCATT  GCATTTCAGC
20601  CTGGGCACAA  AGAGTGAAAC  TCCATCTCAA  AATGAAATAA  AATAACAGAA
20651  ATGTATTTCT  TAACAGTTCT  GGAGGTTGGG  TGGGCAGTCC  CAGATCAGGA
```

<p align="center">FIGURE 3F</p>

```
20701 CACTGACAGA TTCAGTGTCT GATGGGGGCC CACTTTCTGG TGTTACCTGC
20751 TGGCTGTGTT CTCACATGGT GGAAGGAACA TGGCAACTTT CTGGGGCCTT
20801 GTTTTTTAAT TTAAAAAAAA AAAATATTTT CCTGGCCCTT GCCTGCTGAA
20851 GGAACCTCTT TTATAATGGT ACTTAAAAAT TTTTTTTTTT GAGATGGGGG
20901 TCTCACTCTG TCACCCACGC TGAGTGCAGT ATCACAATCT CAGCTCACTG
20951 CAACCTCTGC CTCCCTGGCT TAAGCGATCC TCCCACCTCA GCCTCCTGAG
21001 TACGTGTGAC CATAGGCCCA TGGCACAAAG CCCAGCTAAT TTTTTGTATT
21051 TTTAGTAGAA ATGTGGTTTC ACCATGTTGC ATAGGCTGGT CTCGAACTTC
21101 TGAACTCAAG TGATCTGCCT GCCTTGGCCT CCCAAAGTGC TGGGATTCTA
21151 GGTATGAGCC ACCCTGCTCG GCCTATAATG GCACTTTCCT ATCCCATTGA
21201 TGAGGCTCTA CTCTCATGAC CTAATCATCT CCCAAAGGCC CTAAGGCCTC
21251 CTGATACCAT CACCTTTGGG GTTAGGTTTT AACATATACA TTTTGGGGGG
21301 ACACAGACAT TTTAGACCAT AGCACCTCCA TTGAAAGGAA ACATTTCTGA
21351 CACCTGGCTA TCTCAAAGGG CCCTTTCAGT TCCCCTGCAG GCTGCATTCC
21401 CACATCACCA ACAAGAGCAG CGACACTCAC TCAGGGTTA AATAACTTGT
21451 CCAGAGTCAC AGCAGTAATG AATGACAGAG CTGGGGCTTG AATCCAGGCG
21501 TCCTCCTAGA GCCTGGATTC TGTGTAGTGA GTGAAAGCTG ACTCCTGGGA
21551 GACTTCTGCG TGGTCCTGGT TCTCTCTCCA GACTGCACTG CGCAAGTTTC
21601 TCTTCCTGAT GGTCCCTAGG GTATTACAAA GACAGTGGCC CTGCCTGTCA
21651 GGTGTTTTTA TTACCAGATG AGGTCATGGC CTCAGGAACC CTGTAGGAAG
21701 CTGAGTTCAG AGTCTTTGAG CAGGCTTTAG GGAGGTTCCA GCTTCCCACC
21751 ACCAAGCCCC AGGTGGATTC TTACAGACTC TAGCCTCAGG GTGGGGGGTC
21801 TGGAAGATGA GGTTGCGGGG TGCGATATTC TGCCCAATTC GCCCCTCCTT
21851 GCTCAATCTG TTTCTGCAGG TATTGCTGAC TACAGACCCA AGGATGGAGA
21901 AACCATTGAG CTGAGGCTGG TTAGCTGGTA GCCCCTGAGC TCCCTCATCC
21951 CAGCAGCCTC GCACACTCCC TAGGCTTCTA CCCTCCCTCC TGATGTCCCT
22001 GGAACAGGAA CTCGCCTGAC CCTGCTGCCA CCTCCTGTGC ACTTTGAGCA
22051 ATGCCCCCTG GGATCACCCC AGCCACAAGC CCTTCGAGGG CCCTATACCA
22101 TGGCCCACCT TGGAGCAGAG AGCCAAGCAT CTTCCCTGGG AAGTCTTTCT
22151 GGCCAAGTCT GGCCAGCCTG GCCCTGCAGG TCTCCCATGA AGGCCACCCC
22201 ATGGTCTGAT GGGCATGAAG CATCTCAGAC TCCTTGGCAA AAAACGGAGT
22251 CCGCAGGCCG CAGGTGTTGT GAAGACCACT CGTTCTGTGG TTGGGGTCCT
22301 GCAAGAAGGC CTCCTCAGCC CGGGGGCTAT GGCCCTGACC CCAGCTCTCC
22351 ACTCTGCTGT TAGAGTGGCA GCTCCGAGCT GGTTGTGGCA CAGTAGCTGG
22401 GGAGACCTCA GCAGGGCTGC TCAGTGCCTG CCTCTGACAA AATTAAAGCA
22451 TTGATGGCCT GTGGACCTGC TACAGTGGCC TGGTGCCTCA TACTCCTCAG
22501 GTGCAGGGGC AGGGACAAGA GAAGGGGGAA GTAACCCCAT CAGGGAGGAG
22551 TGGAGGGTGC CTGAGCCGCC ATGTGGGCAT TGGGGGAGTG ATGGGAATGC
22601 CAGCAGTGAT GACGTTGACT ACTGACTGAG CACCCACTAC TATGACTGAG
22651 CACTCACTCG CTAGATACTA TCTTGAACTG CTCTGTGAGG TTGTTGATAT
22701 TTTCATTTTT ATCTGTGCTT TACAAATCAG GAAACTGGGA GCCCGGGCGT
22751 GGTGGCTCAC GCCTGTAATC CCAGCACTTT AGGAGGCCAA GGCAGGTGGA
22801 TCACAAGGTC AGGAGTTTGA GATCAGCCTG GCCAACATGG TGAAACTCCA
22851 TCTTTACTAA AAATACAAAA AATTAGCCAG GCATGGTGTT GCATGCCTGC
22901 ATGCCTGTAA TCCCAGTTAC TTGGGAAGCT GAGGCAGGAG AATTGCTTGA
22951 ACCCTGGAGG CGGAGGTTGT AGTGAGCCGA GATCACGCCA TTGCACTCCA
23001 GCTTGGGCAA GAAGAGAAAC ACTCTCAAAA AAAAAAAAAA ATCAGGAAAC
23051 TGGTGCTCAA AAAGGAAAAG TGACTCACCA AGGTCACAGA CTAGGCAGTG
23101 ATGCTGGGGG AACCTGGCTC AGGGGACACA GACCTGGCCT GGGGCAGCCT
23151 TGCAGCTCCT CCACTAAAAT ACTGAAAATG AGGGGCTTCG ATGATGGTTA
23201 TAATCGTATG GCAGAGCCCC AACTCAACTG GAGCCCTGGG ACCCAGAAGC
23251 TAGGGTCTCA CTCCCTGCTT TTCCACAAGG CACCATTAGG GCATCACCCC
23301 AGGCCTCGGC AGCCACGACG CAGGGATCCT GCCTCTCATT GGTTGGGGGC
23351 TTAGGGGCTC TGGGCTGCCC TCTTGAAGAG GGGGTTCAGC CCAGCGAGGC
23401 ACCCCCTATG CTGCACCCCA CCAAGGTTAG GAAGAGGTCC TGTCCTCAGT
23451 GGGGCCCTCT GATGAACAGC CCATCAGGTC TGCGTCCACA TGCCTTGGAA
23501 GAGATGGTGA CATACTCAAA GTCCTTGAAG CCGCATATTA AACCACCTAG
23551 AGCACCATCT TCAAACATTT AGGGTCTGAG AAGATAGGGG AAGTAAGCAA
23601 TTTAAAACAT TTCTTTATAT TGGGCCAGGT GCAATGGCTC ACGTCTGTAA
23651 TCCCAGCGCT TTGGGAGGAC GAGGATCACC TGAGGTCAGG AGTTCAAGAT
23701 CAGCCTGGCC AACATGGAGA AACCCCATCT CTACTAAAAA TACAAAAATT
23751 AGCTCAGGCG TGGTGATGTG CACCTGTAAT CCTAGCTATT CAGGAGGCTG
23801 AGGCACAAGA ATTGCTTGAG TCAATATTGC ACCACTGCAC TCCAGCCTGG
23851 GCAACAGCGA GACTCTTGTC TCAAAAAAAA AAAAGATAT TTGCTGAAAA
23901 GACCCAGCCT GCCAAACTCA GGGGCAGCCA AGGAGGGTAG TGAAATGGAA
23951 GTTGGAGCTC AGCGCTCCCA CACCTCCACT GCCCTCAGGC CTTCTCTGCC
24001 TCTTTCCCAT CAGTCAGCTG CTTCTGGGCA TGGTCCTGGC AGAGACTTGG
24051 CCTCCTTCCA GTTCAAGCTC CCTCTTAGAT TGTGTCCCAC GCCACTGAGT
24101 CTTTGGGACA CTGGGTCAGA TGTCTAGTCT GGCACAATTG GCAGGAATCC
```

FIGURE 3G

```
24151 CAAGAAACAG TGTGAGTGAG GGGACAGTCG TGTTGAGTGC CCTCCATCTG
24201 GGACTGGGAG GCAGGTCTAT GTCAGGCCTG CATTTAGATC TCTAATGGCT
24251 CCAGACAAGC CCCTTCAGCT CACTAAGCCT GTTTCCTAAC ACAGCTGTGG
24301 GATGGTGCTT TGGTTTACAT AGCACGCGAT ACCATCATAG ATCACATGGG
24351 GAAACTGAGG CCCCAGGAGT GATCTGCTGG CACATGCAGT GACAAGAGGA
24401 GAGGCCCATC TCAGCCTTGC AGCAAGGTTG CCAGAAATCG ATTCTCGCCC
24451 CCATCCCGTA AAGATAGCTG GGATTACAGG TGTGCACCAC CATGCCCAGC
24501 CTAATTTTTG TATTATTAGT AGAGATGGGG TTTCACCATG TTGTCCAGGC
24551 TGGTCATGAA CTCCTGACCT CAAGTGATCC ACCCGCTTTG GCCTCCCAAA
24601 GTGCTGGGAT TACAAGCATG AGCCACAGTG CCTGGCCTGA CCCTGCTCTT
24651 TTGAAAGACC ATTCCCCCAA ATTCTGTGCA CCTGTGTGCC TTTCTTCTCT
24701 CTGCCTCCTC TCAGCTCTGC CCCGCTCTCC TCCCTTCTCC TCTGGCAAAT
24751 CCCACTCATC TCTTGAAGCC CTTCTTCCAG GGGAAGCCCT GATCATGCTG
24801 CTTTCTCCTG TGGGAGGGAT GAAGGACGTG GCCCACGGAG TTTGTTTTGT
24851 TTTGTTTTGA GATGGAGTTT TGCTCATGTT GCCCAGGCTG GGGTACAATG
24901 GTACGATCTC AGCTCACTGC AACCTCTACG TCCCGGGTTC AAGCGGTTCT
24951 CCTGCCTTAG CCTCCCCAGT AGCTGGGATT ACTGGCATGA ACCACCACAC
25001 CTGGCTAATT TTGTGTTTTT AGTAGAGATG GGGTTTCTTC ATGTTGGTCA
25051 GGCTGGTCTC GAACTCCCAA CCTCAGGTGA TCTGCCTGCC TCGGCCTCCC
25101 AAAGTACTGG GATTACAGGG TTGAGCCACT GTGCCTGGCC CAGGCCCACG
25151 GAGTTTTAAG AGGCTTCCTG TGGCAGTGGC ATCCAGACGG AGTGCAGAAA
25201 CTCAAAGTTG AAGGCCAGAA GCTCAGGGAA GGGGGACTGT GAGTTGAGGA
25251 GTCTCTTGGC TGCCAGGGCC AGAAACCGAA CTCCAAGCCT CTCCACAACA
25301 GCGGGTGTAG AGCATGTAGA ATCAGAGAGG AGGCTGAGCC ATGCAGCCCC
25351 GAGAAGAGGG GAATGCCACT GAGCCACAGA GACCCAGTGC CACTGCCAGG
25401 TGTCTCTGCC TCCACTTCCC ATGACCCGGC CTGTCTCTGT ATGCAGGCTT
25451 CACCCTCTCT CGTTGTACAT TGTACACATT CTAGGTGACA CCAGCAGCTT
25501 CTGATTCTCA TCTCCCATAA CATCAGCCCC CCAGAGAGGG GACAACTGCT
25551 GAGCTGATAA CATAATAGAT GCCCCTTTCC TGGAGCGGCC AT GGTCATGGTC
25601 AGCGTGGAGA GGATGAAGCC TGAGCAGGCA GGATCGGGGG TCTAGAGGGG
25651 AAGGAGGTGG AAGTTGAGAT CACAGACCTG TGGTCAGGTG GCCTGGGAAG
25701 GGTTTGACGA GTGTCGGCCC AAAGAGCTTG GAAGGGATTT TGCTGCTGTG
25751 GGTGAGCACT GCCTCTCCCC TTAGGGACAA CAGCCACCTC TTCTCTCCCC
25801 ATTTGCCTTT CCCTTCTGTA GATATGAAAC ACAGGCCTCC TTGTCAGGCC
25851 CCTACTTAAC CTCCGTGATG GGGAAAGCGG CCGGAGAAAG GGAGTTCTGG
25901 CAGCTTCTCC GAGACCCCAA CACCCCACTG TTGCAAGGTG AGTCATGGCC
25951 TGACACTCTG GATGTGTCCC CTACCCCAAG CTTACTCAGC CAAGAGGCTT
26001 CATCAACTCA CCCCAGCTTT CCCTAGCACC CTCCTGGGCC ACACCTTCAC
26051 AAAATCACTG ATGCTCAAAG TTGGATATAA TATATTGAAC TGAAGCCTTA
26101 GCATTTTTAT GCAAGTTACT GTGGAAATTC TAGGAAACCA GACAGATTAC
26151 AAAAAAAAAA AAAAACTAGA AGAAAATTAA CATCACCTAG GATATACTAC
26201 CTAGGAATAA CGTCTTTTAT TTTGAGATGG AGTTTCGCTC TTGTTGCCCA
26251 GGCTGGAGTG CAGCGGTATG ATCTCGGCTC GCTGCAACCT CCGCCTCCTG
26301 GGTTCATGTG ATTCTTCCAC CTCGGCCTTC CTAGAGCCCA AGTGGTGTGC
26351 CTGCCTCTGC CTCCCAAAGT TCTGGGATTA CAGGCATGAG CCACCGCACC
26401 CAGCCAAAAT TACTTAACTT TTCTTCTAGA TACTTTTTAA AAATATGGCA
26451 GTAAGTTTTT CATAAAAAAT GGAGCCATGC TATCCAGTGG AAATTTAATG
26501 TTGCCCACAT GTATAACTTA AAAATTTCAT ATATGTGTAT ACATATATAT
26551 GAAATATATA TATACAGACA CACATATATA TGTATACATA TATATACACA
26601 TATATATGTA TACATATATA CACACATATA TGTATACATA TATATACACA
26651 CATATACACA TATATACACA CACATACATA TATACACATA CATATATACA
26701 CACATATATA CACACATGCA CACATATATA TGTATACATA TATACACACA
26751 TGTATACGTA TATATACACA CATATATACA CACATATATA TACACACATA
26801 TACACACATA CACACACATA TATACACACA TATATACACA CATATATACA
26851 CACATATATA TGTATACATA TATATACACA CATATATACA CATACACACA
26901 TACATATATA CACATATACA CATATACACA CACATATACA CACATGTATA
26951 CATATATATA CACACATGTA TACATATGTA TACACACACA TATATGTATA
27001 CATATATACA CACATACATA TGTGTACATA TATACACACA TACATATGTA
27051 TACATATATA CACACAT
(SEQ ID NO:5)
```

FIGURE 3H

Isoform 1:

FEATURES:

Exon:	2031-2094
Intron:	2095-5569
Exon:	5570-5762
Intron:	5763-7571
Exon:	7572-7741
Intron:	7742-10000
Exon:	10001-10173
Intron:	10174-10298
Exon:	10299-10485
Intron:	10486-12027
Exon:	12028-12193
Intron:	12194-25821
Exon:	25822-25939

Allelic Variants (SNPs):

DNA Position	Major	Minor	Domain	Protein Position	Major	Minor
921	C	T	Beyond ORF(5')			
1781	C	T	Beyond ORF(5')			
1850	G	A	Beyond ORF(5')			
2839	A	G	Intron			
3730	G	A	Intron			
6631	G	A	Intron			
6945	–	A	Intron			
6952	A	T	Intron			
7457	G	A	Intron			
7830	T	A	Intron			
8089	T	C	Intron			
8551	C	T	Intron			
9269	G	C	Intron			
9362	C	T	Intron			
9782	G	T	Intron			
11493	G	A T	Intron			
12260	A	G	Intron			
13086	T	C	Intron			
13183	T	C	Intron			
21240	C	G	Intron			
21695	A	G	Intron			
22058	C	T	Intron			
22233	C	A	Intron			
22245	C	–	Intron			
22375	C	T	Intron			
23042	A	– T	Intron			
23344	T	C	Intron			
23873	A	–	Intron			
24764	G	T	Intron			
24939	T	C	Intron			
24945	G	A	Intron			
25092	C	T	Intron			
25428	T	G	Intron			
25513	C	T	Intron			
25684	C	T	Intron			
26165	A	–	Beyond ORF(3')			

Context:
DNA
Position

921   TTGGAGATATTTTAAGGTCATAGTGTCTTCACAAATTGAGCTGAAAGGGAACTGTTAGGA
TGATCTTGCCTAACCCTCTCATCTCACACAGGAAGAACTATTTTAAACTCGAGAGGTTAA
GTGACCTGGCCAAAGTCACACAGCCACCACTAGTTAACTCGTATACATTGATTCTCCTGT
GGGGCTGGGCAGATGAGGAATCTTTTGTTCTCTTCCCTGTTTGCAGAGATTTTTTTTGAG
GTTACTTTCCGAGTTCTGGCAAGTACCCCTGCTTCTGGTAGCTTTGTGTCTCGATTCAAT
[C,T]

FIGURE 3I

```
          TCATTCTTTTTATTTTATTTTATTTTTGAGACAGGGTCTCACTTTGTCACCCAAGCTGGA
          GTGCAGTGGTGTAATCTTGGCTCACTGTAGCCTCCACCTCTTGGGTTCAAGCGATCCTCC
          TGCCTCAGCCCCCCAAGTAGCTGGGATTACAGACGTCTGCCACCACGCCAGGCTAATTTA
          TGGTTTTTTGTATGTGTTTTTTGTGTTTTTGTAGAGACAGTGTTTCCCCATGTTGCCCAG
          GCTGGTCTCCAACTCCTGAGCTCAAGTGATCTGCCCGCCTCAGCCTTTCAAAGTGCTAGG

1781      ACAGCTGACTCCAGCAATGCTGCTCACGTGACCACTGCAGCTGCAGCTCCCGTTCCACTC
          CTTGTCCTGGGCTAGGTGGGCACTACCAGGGGCTCCTTTGGTAAGGAGTACCGGGTAGGC
          ACCCGGTCCTGCCAATCCACCACTGGAACAGCTGGGGGGACAGCAGACAGGCACGGTCGG
          ACAGACTTGACAGATCAGGCATCAGGCCCTCTGCGCTGGTCCCGGGCTCTTTAAGCAGGA
          ACGTGAATGGCCTCAAGATGTCTCACATGGTCCCACTAGCCCTCCTCCTCCCTTTGTTCC
          [C,T]
          TACCTCCAGGAGGGCTGCTCTGCCCTTCCTTCCTCTGTTCTTTGGCCTTATGTTCCCCGC
          CACCCACAGGCCTTCCCCCGCCCCACCCCTCTGCAGACTTAGCCGTGCATTGCAGGCATGG
          AGGATTAATCAGTGACAGGAAGCTGCGTCTCTCGGAGCGGTGACCAGCTGTGGTCAGGAG
          AGCCTCAGCAGGGCCAGCCCCAGGAGTCTTTCCCGATTCTTGCTCACTGCTCACCCACCT
          GCTGCTGCCATGAGGCACCTTGGGGCCTTCCTCTTCCTTCTGGGGGTCCTGGGGGCCCTC

1850      GGCTAGGTGGGCACTACCAGGGGCTCCTTTGGTAAGGAGTACCGGGTAGGCACCCGGTCC
          TGCCAATCCACCACTGGAACAGCTGGGGGGACAGCAGACAGGCACGGTCGGACAGACTTG
          ACAGATCAGGCATCAGGCCCTCTGCGCTGGTCCCGGGCTCTTTAAGCAGGAACGTGAATG
          GCCTCAAGATGTCTCACATGGTCCCACTAGCCCTCCTCCTCCCTTTGTTCCCTACCTCCA
          GGAGGGCTGCTCTGCCCTTCCTTCCTCTGTTCTTTGGCCTTATGTTCCCCGCCACCACAG
          [G,A]
          CCTTCCCCCGCCCCACCCCTCTGCAGACTTAGCCGTGCATTGCAGGCATGGAGGATTAAT
          CAGTGACAGGAAGCTGCGTCTCTCGGAGCGGTGACCAGCTGTGGTCAGGAGAGCCTCAGC
          AGGGCCAGCCCCAGGAGTCTTTCCCGATTCTTGCTCACTGCTCACCCACCTGCTGCTGCC
          ATGAGGCACCTTGGGGCCTTCCTCTTCCTTCTGGGGGTCCTGGGGGCCCTCACTGAGATG
          TGTGGTGAGTAACTCGCCTCTATCCTGTGCCTCTTTCCTCCTGGGTCCTTAGTGGGGTGG

2839      AACATAGGGAGACCCCATCTCTACAAAAAATAAAAAAAATTAAAAAATAGCTGGGCATGG
          GGAAGACTTTCTGAAGACCAAGAGGACACATGGGAGCTGAAACTCGAAGGAAGAAAAGGA
          GCTGGCAGGAAAGGAGTGGGGGACACACATTCTAGGCAGCAGGAAGTGAGCCTTCGGAGG
          TCCTGCCTGCTCCAGCTCTGTGCCCCAAGGGGTCTCTTGGAGCACAGTCTCCTGGGACCT
          GTCTATGAGTCTGAGCTTAGAGGCTCAGGGCTGCTCCTTCAGACAGGAGGCAGAAGGCAG
          [A,G]
          CTTTGGGAACTTTGGGCCGCCCACGCGCCCTTTTCTCCTCCTCTGCACCTAGGATTACGTT
          GAGCAATACACTTTCACCCCCATGGTCTCTTGAGACCCTGGGGAAACCCTGAGAGGTGGG
          TGCAGTCATGTCCAGGTGTCAAGTGAAGAAGTCGAGGGTTGGAGGGGCTGAGTGACCCAC
          TCAGGGTGCTCCACCTTTTCCAGAGCTTTGCTGAACTTAGTTTTTTAGAACTTGAAGCCTC
          GTTTGTTTTCGTTTTGTTTTTTGTTGAGAGAGGTTCTCCCTCTGTTGCCCAGGCTGGAGT

3730      GACACCTCAGGTCTGGGCCCAGGAACCCCAGCTCTTGGTTCATGTCCGGACAGTCCCCAG
          GGGAGTTCTGGGTTCAACCAGCAAGAGCTCTTCCTCCTGGCTGATCTGGTCCTCAGCCTT
          GGACAGTTAGTCCATTAACCTGACCCCACAGGAGCCCCAATCCCTTGGGGTCTGGGGAAT
          CTTGAACTGGGGTTTGGGGTGCAAATATCTGCACTGAGTCACTTAATTGCACCCAGCCTC
          ATTCCTTTATCTGTAAAGTGGGCTAAGAATGCTCCCCTGCCTTCCTCCTCGGTGTAGTAC
          [G,A]
          AGGAAGGATCCCATGACACCTGCTCTCCCAGTTTAAAGCTCTATATGTATGTTGTGAAAT
          TGACAGGGATCGCTGCACAAACGCTAATGCAAAGTGGGCTCCTGTGCTTCCTTTTCTCTT
          TCTTCTTCTTTTTTTTTTTTTTAATTTTCTTCTAGAGATGAGGTCTCACTATATTGCCCA
          GGGTTGGTTTCAAACTCCTAGGGTCAAGCGATCCTCCCACCTTGGCCTCCCAAACTGCTG
          GTATTACAGGCGTGAGCCACTCTGTCTGGCTCCTATGCTTGTGAATGTCAACAGCAATCA

6631      TGAATAGCTGGGATTACAGGCGTGTGCCACCATGCCCAGCTGATTTTTGTATTTTTGGTA
          GAGATAGGGTTTCACCATGTTGGCCAGGCTGGTCTTGAACTCCTGACCTCAGGTGATCCG
          CCTGTCTTGGCTTCCCAAAGTGTTGGGATTATAGGCATGAGCCACTGCACCAATCCAAAA
          GCAGCATCTTTGTGCTCCCTTTTCAAGAGGCATCACAGAGAGGCCTGTTTTGGGGTTTGA
          ATGAGAGGCGAAGAATCAGCCATGGAGTGCCTCTTTCTCAGACTCCCTCTTGAGAAGTGG
          [G,A]
          TGCAGGGGTGGAGAGAAAAGAAGACTAGGCATAGTGGCTCATACCTGTAATCCCAACATT
          TTGGGAGGCTGAGGCAGGAAGATTGCTTGAGCTCAGGAGTTTGAGACCAGCCTAGGCAAC
          ATAGTGAGACCACATCTCTTAAAAAAAAAGAAAAAGAAAAAAAAATGAGCCAGGTGTAGTGA
          CTCATGCCTGTGGTCCCCACTTCTCCGGAGGCAAAGGTGGGAGGATCTTTTGAGGCTGAG
          AAATCGAGGCTACAGTGAGCCATGGTGGCACCACTGCACTCCAGCCTGGGAGACAGAGAG

6945      AGAAAAGAAGACTAGGCATAGTGGCTCATACCTGTAATCCCAACATTTTGGGAGGCTGAG
          GCAGGAAGATTGCTTGAGCTCAGGAGTTTGAGACCAGCCTAGGCAACATAGTGAGACCAC
          ATCTCTTAAAAAAAAAGAAAAAGAAAAAAAAATGAGCCAGGTGTAGTGACTCATGCCTGTGG
```

FIGURE 3J

TCCCCACTTCTCCGGAGGCAAAGGTGGGAGGATCTTTTGAGGCTGAGAAATCGAGGCTAC
AGTGAGCCATGGTGGCACCACTGCACTCCAGCCTGGGAGACAGAGAGACCCTATCTCAGT
[-,A]
AAAAAAAAAAATAAAAATATGGCTGGGTGTGGTGGCTCACGCCTGTAATCCCAGCACTTT
GGGAGGCCAAGGTAGGTAGATCACATGAGGTTAGGAGTTCGAAACCAGTCTGGCCAACAT
AGTGAAACCCTGTCTCTACTGAAAATACAAAAAATTAGCCAAGGGTGGTGGTGGGCAACT
GTAATCCCAGCTACTTGGGAGGCCGAGGCAGAAGAATCGCTTGAACTCGGGAGGCGGAGG
TTGCAGTGAGCTGAGAACATGC

6952   AAGACTAGGCATAGTGGCTCATACCTGTAATCCCAACATTTTGGGAGGCTGAGGCAGGAA
          GATTGCTTGAGCTCAGGAGTTTGAGACCAGCCTAGGCAACATAGTGAGACCACATCTCTT
          AAAAAAAAGAAAAAGAAAAAAAAATGAGCCAGGTGTAGTGACTCATGCCTGTGGTCCCCAC
          TTCTCCGGAGGCAAAGGTGGGAGGATCTTTTGAGGCTGAGAAATCGAGGCTACAGTGAGC
          CATGGTGGCACCACTGCACTCCAGCCTGGGAGACAGAGAGACCCTATCTCAGTAAAAAAA
          [A,T]
          AAAATAAAAATATGGCTGGGTGTGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGC
          CAAGGTAGGTAGATCACATGAGGTTAGGAGTTCGAAACCAGTCTGGCCAACATAGTGAAA
          CCCTGTCTCTACTGAAAATACAAAAAATTAGCCAAGGGTGGTGGTGGGCAACTGTAATCC
          CAGCTACTTGGGAGGCCGAGGCAGAAGAATCGCTTGAACTCGGGAGGCGGAGGTTGCAGT
          GAGCTGAGAACATGCCACTGCACTCCAGCCTGGGCAACAAGAGCGAAACTCTGTCTCAAA

7457   AAGAATCGCTTGAACTCGGGAGGCGGAGGTTGCAGTGAGCTGAGAACATGCCACTGCACT
          CCAGCCTGGGCAACAAGAGCGAAACTCTGTCTCAAAGAAAATAAATAAATAAAATAAAAA
          AATAAAAAAGGAGGGGGCATATGGGTGAAGTATGGACAAAATAGTGGGGCAGGCACAGAT
          GATCTGGACACAGGAGCCCTTGGAGTTTATTCTTGAATCTAACTGTTCATCTTTATTAAA
          TATTTGTGGCATACACCTCACAACAACATAGCCAACACACCTCCTTTTGGAGCTTTTATC
          [G,A]
          AAGTTTCCCACTGTTAAGATTTTTTTCCCGCTTTGTGATGCGGGTGGGGTGGGTGCTGTAA·
          GCAGGCTTACGGGGTGGCAGTTTCTCACAAAGGCATTAACTGGCCTTGTCCTAGGTCTGC
          CTTCAGCGAGGATGACGGTGACTGCCAGGGCAAGCCTTCCATGGGCCAGCTGGCCCTCTA
          CCTGCTCGCTCTCAGAGCCAACTGTGAGTTTGTCAGGGGCCACAAGGGGGACAGGCTGGT
          CTCACAGCTCAAATGGTTCCTGGAGGATGAGAAGAGAGCCATTGGTGAGCAGACACCATC

7830   GGTGGCAGTTTCTCACAAAGGCATTAACTGGCCTTGTCCTAGGTCTGCCTTCAGCGAGGA
          TGACGGTGACTGCCAGGGCAAGCCTTCCATGGGCCAGCTGGCCCTCTACCTGCTCGCTCT
          CAGAGCCAACTGTGAGTTTGTCAGGGGCCACAAGGGGGACAGGCTGGTCTCACAGCTCAA
          ATGGTTCCTGGAGGATGAGAAGAGAGCCATTGGTGAGCAGACACCATCCGCTGGGGGTGG
          GGAGCAGCTGGGAGGGCTCATCAGATGATATTCTCCAATGAGAATCAGAACTTTGGGTTT
          [T,A]
          CTCCCCAGGCGTCTTTCCCACCATCCATTCTGCCCATCTCACTGCCTACGTAGAGGCTCG
          AACCTGTCCCCATAGCCATCCTTGACCCAGCTTTTCCCGCGCTGCACACATACTATTGAC
          AGGTGTGTTTCGTGGTTTTTTGTTTTTTGTTTGTTTGTTTGTTTTGAGTTGGAGGTTTGC
          TCTTGCTGCCCAGGCTGGAGTACAATGGCGCAATCTCAGCTCACCGCAATCTCTGCCTCC
          TGGGTTCAAGCAATTCTCTTGCCTCAGCCTCCTGAGTAGCTGGGATTACAGGCATGCGCC

8089   ATCAGATGATATTCTCCAATGAGAATCAGAACTTTGGGTTTTCTCCCCAGGCGTCTTTCC
          CACCATCCATTCTGCCCATCTCACTGCCTACGTAGAGGCTCGAACCTGTCCCCATAGCCA
          TCCTTGACCCAGCTTTTCCCGCGCTGCACACATACTATTGACAGGTGTGTTTCGTGGTTT
          TTTGTTTTTTGTTTGTTTGTTTGTTTTGAGTTGGAGGTTTGCTCTTGCTGCCCAGGCTGG
          AGTACAATGGCGCAATCTCAGCTCACCGCAATCTCTGCCTCCTGGGTTCAAGCAATTCTC
          [T,C]
          TGCCTCAGCCTCCTGAGTAGCTGGGATTACAGGCATGCGCCACCACACCCAGCTAATTTT
          GTATTTTTAGTAGACGTGGGGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCTGACC
          TCAGGTGATCCGCTTGCCTTAGCCTCCGAAAGTGCTGGGATTACAGGCATGAGCCACTGC
          GTTAGGCCCACTGACAAGCCTTGTATTGGCTAGCCACCAAGATTGACTTGATTATCCACC
          TTCGGGACAACTGGACAGCCTGCTTATGACTTACGCCATAGTCTGTCTCTACTAGCTCTC

8551   TACAGGCATGAGCCACTGCGTTAGGCCCACTGACAAGCCTTGTATTGGCTAGCCACCAAG
          ATTGACTTGATTATCCACCTTCGGGACAACTGGACAGCCTGCTTATGACTTACGCCATAG
          TCTGTCTCTACTAGCTCTCCTGCCCTGACTTGACCCAGCATACAACAGCCAGAGCCAGCC
          TTTTCAATATAAACCTGATCTTGCTGGCACTGCTTAAACCCTGCAGGGGCCTCGCACTGC
          TCCATGGCCCAGCCTGTCTACCCTTACCTTCTGCCCAGGCTCTGCTCATCCATTCTCTGC
          [C,T]
          TCCCACACACCTGCCCTCTGTGGGCTCCAGCCATACCATCTCTCAACTCATAAGCCAGTT
          TTTTCATACAGGCTCCCTCCATCTGGACTGGCTTCCCTGCGTGCAGTTCACTCCTGCTCT
          ACCTTTGGCTCTGCCTCCACCCATCCTCAGCCGTCTCCAGCATTACCTCCTTGGAGAATC
          CTGCCTTGACTTCCCAGCCACCCAAATATCACTACTTGGTCTGCATTCTCGTTGCAATTG
          CAGTCGCATGAGCAATTGCTGTGGTTGAGGCCCGAACTGCGCAAGTGCCTGTCTGCCATG

FIGURE 3K

9269   AGGCCAGGGTCCCAGGTGCTGGCGGGGCTGGCTGCTGGGTGGGGGCAGAGAGGCAACCCC
TCTGTTTTTTTCCCTCTCAGGGCATGATCACAAGGGCCACCCCCACACTAGCTACTACCA
GTATGGCCTGGGCATTCTGGCCCTGTGTCTCCACCAGAAGCGGGTCCATGACAGCGTGGT
GGACAAACTTCTGTATGCTGTGGAACCTTTCCACCAGGGCCACCATTCTGTGGGTGAGTA
GGTCAGACCGTGCCAAGGCCAGGCTGGCACTCCCTCAGTCCCCAGGTCTGCACTGATGAC
[G,C]
TCCATACCCTGGCCCCCACACTCACCTTTCCTTGGGGCTCCTCCGAATCAAGTCCTTTAG
GGACGAATTGGCGAGGGCTCATGGGTGATGCTCCAGCTGTGAGCCAGCTTTGGAGCTGGT
AGGTGGATCTCTTGAGGCCAGGAGTTCAAGACAACGTGGTGAAACCCCATCTCTACTAAA
AATAAAAAAGTTAGCCGGGCATGGTGGCACATGCCTGTAGTCCCAGCTACTCGGGAGGCT
GAGGCAGGAGAATCACTTGAACCTGGGAGGCGGAGGCTGCAGTGAGTGGAGATCGCACCA

9362   GGGCCACCCCCACACTAGCTACTACCAGTATGGCCTGGGCATTCTGGCCCTGTGTCTCCA
CCAGAAGCGGGTCCATGACAGCGTGGTGGACAAACTTCTGTATGCTGTGGAACCTTTCCA
CCAGGGCCACCATTCTGTGGGTGAGTAGGTCAGACCGTGCCAAGGCCAGGCTGGCACTCC
CTCAGTCCCCAGGTCTGCACTGATGACGTCCATACCCTGGCCCCCACACTCACCTTTCCT
TGGGGCTCCTCCGAATCAAGTCCTTTAGGGACGAATTGGCGAGGGCTCATGGGTGATGCT
[C,T]
CAGCTGTGAGCCAGCTTTGGAGCTGGTAGGTGGATCTCTTGAGGCCAGGAGTTCAAGACA
ACGTGGTGAAACCCCATCTCTACTAAAAAATAAAAAAGTTAGCCGGGCATGGTGGCACATG
CCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATCACTTGAACCTGGGAGGCGG
AGGCTGCAGTGAGTGGAGATCGCACCACTGCCCTCCAGCCTGGGCAACAGAGTGAGTGAG
ACTCTGTCTCAAAAAATAAAAAATAAAATAAAACTCCCCTAGTGATTCCAATGTGCAGCT

9782   GCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATCACTTGAACCTGGGAGGCG
GAGGCTGCAGTGAGTGGAGATCGCACCACTGCCCTCCAGCCTGGGCAACAGAGTGAGTGA
GACTCTGTCTCAAAAAATAAAAAATAAAATAAAACTCCCCTAGTGATTCCAATGTGCAGC
TAAGTTTGGAAATAGGTGGTATGGGGTCAAGTCCTCTTGGGCCTCCCTCCTCCAGTCCTT
CTCCCTAACCTCTAGCCCTCAAGTTGCAGAGTGATCAGCCAAACCAGTTTGCCCAGAAAT
[G,T]
AGCAGTTTCCTGGGACACAGGATTTTCAGAGTCCAGACAAGGAAAGTCTTGGGCAGACCA
GGTTGAGTTGGTGCCCTTAGCTGATCTGACCATGTTGCCCTTCTTCTCCAAGCCCTCCTG
TGGTTGTCCATAGCTACAAGGGCCTGACCCTCAAGCCCCTGCCTGTCCTGGCCCCTTTGG
CTCTCCAGCTCATTGCATGTTCTGTCCCCCACTTCAAGACACAGCAGCCATGGCAGGCTT
GGCATTCACCTGTCTGAAGCGCTCAAACTTCAACCCTGGTCGGAGACAACGGATCACCAT

11493   AAAAAAAAAAAATGGAGAAGAAGGAAGCTGGACATGGTGGCTCGTGCTTATAATCCTAGCA
CTCTGGGAAGCTGAGGCAGATGGATTGCCTGAGCCCAGGAGTTTGAGACCAGCCTGGGCA
ACATGGTGAAACCCTGTCTTTACTAAAATACGAAAGATTAGCCAGGCATGGTGGTAGACA
CCTATAATCCCAGCTACTAGGGAGGCTGAGCCACAAGAATCACTTGAACCTGGGAGACAG
AGGTTGCAGTGAGCCGAGATCGCGCCATTGCACTCCAGCCTGGGCGACAGTGTGAGACTC
[G,A,T]
GTCTCCAGAAAAAACAAGAATGGATAGAGTGGAGCCAAGAAGAGGCAGGAAGAACAAAGA
CACAGAGGTGCACAGAGTTTGGGGGAATTTTGAGGAATGGTCTTGCAAAAGAGTGGGATC
TGGGAGAATGAGTGGGAGTGGAAAGCAGATGAATGAAGAGAAGGTGAGCGCATCAGGGTA
ACAGAGATGCGTTGTGAACAAATGCATGTTCTAGGAAGAGCCCTCTGGAGTGCTAGGTGC
CAGAGAGGTGGGAGGAAGGATACTGGAAGCAGAGAAACCAGTGAGGGGCCTGATCTTGGG

12260   ACAAGAAGACAAATAATCCAGGCTCTCTGTCCTCACACCAGCTGCCCGCCCCTTTCTTCC
TGGCACAGTCATGTTGGAACCAGCTGCTGAGACCATTCCTCAGACCCAAGAGATCATCAG
TGTCACGCTGCAGGTGCTTAGTCTCTTGCCGCCGTACAGACAGTCCATCTCTGTTCTGGC
CGGGTCCACCGTGGAAGATGTCCTGAAGAAGGCCCATGAGTTAGGAGGATTCACGTGAGA
CTCCCACCTCCCAGTCCTCACCCCACCCAACCTCACATGCCTGATAACAGGGTCACAGAA
[A,G]
AGACGGGGAACAGAGGAGAGGGTTCCCTCGGGAGAGACACTGGCCCTGCTTCTGCTTCTA
CCTGCTCAGCTCCTTTCTTGCCCACGGTGTTATGGAAACAGGGAGCCATAGGCCAGCATT
GTCACTGAGAGAGCAGGCTTTGGAGGCAGAGCCCCCCAGTTGGAATCCCAACTCTAACCA
GCTAGGTTCCAGGTAGGCACCCACAATTCACCGAGGAGAACAGTTGTGCCCCTTCCCTGC
AGGGCCAGTGTGAAGAGTCCAGGAGTTAGTACACATAGAGATAGTGGCATGTGCTTTTTA

13086   GGCACGTGCCACACAACCCTGGGTAATTTTTTTTTTTTTTTTTTGAGATAGGGTCTCTG
TCTGTTGCCCAGGCTGGTCTCAAATTCCTGGCCTCAAACCATCCTCACACCTGAGGCGCT
CAAAATATTGGGATTATAGGTGCGAGCCATCATGCTCAGCCAGAATAATAACTGGTTTTT
TTTGTTTTTTTTTTGAGACAGAGTCTCACTCTATTACCCAGGCTCTGGAGGCCCAACTCG
TGTTTGTGTATTTGTTTATTTTTATTTATTTATTTATTTCGAGACAGAGCCTCTCTCTTT
[T,C]
ACCTAGGCTGGAGTGCAGTGGCGCAATCTCGGCTCACTGCAACCTCCGTCTCCTGGGTTC
AAGTGATTGTCCTGCCTCAGCCTCCTGAGTAGCTGGTGCTACAGGCGCGTGCCACCATGC
CCAGCTAATTTTTGTATTTTTAGTAGAGACAGGGTTTTACTATGTTGGCCAGCTGGTTTC

FIGURE 3L

```
              TAACTCCTGAACTCGGGTGATCTGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGC
              ATGGGCCTCCGTGCCCGGCCATGTATTTATTTAGGCAAGGTCTCTCTCTGTTATCCAGGC

13183         ACCATCCTCACACCTGAGGCGCTCAAAATATTGGGATTATAGGTGCGAGCCATCATGCTC
              AGCCAGAATAATAACTGGTTTTTTTTTGTTTTTTTTTTGAGACAGAGTCTCACTCTATTAC
              CCAGGCTCTGGAGGCCCAACTCGTGTTTGTGTATTTGTTTATTTTTATTTATTTATTTAT
              TTCGAGACAGAGCCTCTCTCTTTCACCTAGGCTGGAGTGCAGTGGCGCAATCTCGGCTCA
              CTGCAACCTCCGTCTCCTGGGTTCAAGTGATTGTCCTGCCTCAGCCTCCTGAGTAGCTGG
              [T,C]
              GCTACAGGCGCGTGCCACCATGCCCAGCTAATTTTTGTATTTTTAGTAGAGACAGGGTTT
              TACTATGTTGGCCAGCTGGTTTCTAACTCCTGAACTCGGGTGATCTGCCTGCCTCGGCCT
              CCCAAAGTGCTGGGATTACAGGCATGGGCCTCCGTGCCCGGCCATGTATTTATTTAGGCA
              AGGTCTCTCTCTGTTATCCAGGCTGAAGTGCAGTGGCACATTCATAGCTCACTGCAGCCT
              CAAATTATCCAAGTAACAGGGACTACAGGCATGCACCACCACACCCATCTACTTTTTTTT

21240         TCAGCTCACTGCAACCTCTGCCTCCCTGGCTTAAGCGATCCTCCCACCTCAGCCTCCTGA
              GTACGTGTGACCATAGGCCCATGGCACAAAGCCCAGCTAATTTTTTGTATTTTTAGTAGA
              AATGTGGTTTCACCATGTTGCATAGGCTGGTCTCGAACTTCTGAACTCAAGTGATCTGCC
              TGCCTTGGCCTCCCAAAGTGCTGGGATTCTAGGTATGAGCCACCCTGCTCGGCCTATAAT
              GGCACTTTCCTATCCCATTGATGAGGCTCTACTCTCATGACCTAATCATCTCCCAAAGGC
              [C,G]
              CTAAGGCCTCCTGATACCATCACCTTTGGGGTTAGGTTTTAACATATACATTTTGGGGGG
              ACACAGACATTTTAGACCATAGCACCTCCATTGAAAGGAAACATTTCTGACACCTGGCTA
              TCTCAAAGGGCCCTTTCAGTTCCCCTGCAGGCTGCATTCCCACATCACCAACAAGAGCAG
              CGACACTCACTCAGAGGTTAAATAACTTGTCCAGAGTCACAGCAGTAATGAATGACAGAG
              CTGGGGCTTGAATCCAGGCGTCCTCCTAGAGCCTGGATTCTGTGTAGTGAGTGAAAGCTG

21695         CATTCCCACATCACCAACAAGAGCAGCGACACTCACTCAGAGGTTAAATAACTTGTCCAG
              AGTCACAGCAGTAATGAATGACAGAGCTGGGGCTTGAATCCAGGCGTCCTCCTAGAGCCT
              GGATTCTGTGTAGTGAGTGAAAGCTGACTCCTGGGAGACTTCTGCGTGGTCCTGGTTCTC
              TCTCCAGACTGCACTGCGCAAGTTTCTCTTCCTGATGGTCCCTAGGGTATTACAAAGACA
              GTGGCCCTGCCTGTCAGGTGTTTTTATTACCAGATGAGGTCATGGCCTCAGGAACCCTGT
              [A,G]
              GGAAGCTGAGTTCAGAGTCTTTGAGCAGGCTTTAGGGAGGTTCCAGCTTCCCACCACCAA
              GCCCCAGGTGGATTCTTACAGACTCTAGCCTCAGGGTGGGGGGTCTGGAAGATGAGGTTG
              CGGGGTGCGATATTCTGCCCAATTCGCCCCTCCTTGCTCAATCTGTTTCTGCAGGTATTG
              CTGACTACAGACCCAAGGATGGAGAAACCATTGAGCTGAGGCTGGTTAGCTGGTAGCCCC
              TGAGCTCCCTCATCCCAGCAGCCTCGCACACTCCCTAGGCTTCTACCCTCCCTCCTGATG

22058         CCCAGGTGGATTCTTACAGACTCTAGCCTCAGGGTGGGGGGTCTGGAAGATGAGGTTGCG
              GGGTGCGATATTCTGCCCAATTCGCCCCTCCTTGCTCAATCTGTTTCTGCAGGTATTGCT
              GACTACAGACCCAAGGATGGAGAAACCATTGAGCTGAGGCTGGTTAGCTGGTAGCCCCTG
              AGCTCCCTCATCCCAGCAGCCTCGCACACTCCCTAGGCTTCTACCCTCCCTCCTGATGTC
              CCTGGAACAGGAACTCGCCTGACCCTGCTGCCACCTCCTGTGCACTTTGAGCAATGCCCC
              [C,T]
              TGGGATCACCCCAGCCACAAGCCCTTCGAGGGCCCTATACCATGGCCCACCTTGGAGCAG
              AGAGCCAAGCATCTTCCCTGGGAAGTCTTTCTGGCCAAGTCTGGCCAGCCTGGCCCTGCA
              GGTCTCCCATGAAGGCCACCCCATGGTCTGATGGGCATGAAGCATCTCAGACTCCTTGGC
              AAAAAACGGAGTCCGCAGGCCGCAGGTGTTGTGAAGACCACTCGTTCTGTGGTTGGGGTC
              CTGCAAGAAGGCCTCCTCAGCCCGGGGGCTATGGCCCTGACCCCAGCTCTCCACTCTGCT

22233         CCCTGAGCTCCCTCATCCCAGCAGCCTCGCACACTCCCTAGGCTTCTACCCTCCCTCCTG
              ATGTCCCTGGAACAGGAACTCGCCTGACCCTGCTGCCACCTCCTGTGCACTTTGAGCAAT
              GCCCCCTGGGATCACCCCAGCCACAAGCCCTTCGAGGGCCCTATACCATGGCCCACCTTG
              GAGCAGAGAGCCAAGCATCTTCCCTGGGAAGTCTTTCTGGCCAAGTCTGGCCAGCCTGGC
              CCTGCAGGTCTCCCATGAAGGCCACCCCATGGTCTGATGGGCATGAAGCATCTCAGACTC
              [C,A]
              TTGGCAAAAAACGGAGTCCGCAGGCCGCAGGTGTTGTGAAGACCACTCGTTCTGTGGTTG
              GGGTCCTGCAAGAAGGCCTCCTCAGCCCGGGGGCTATGGCCCTGACCCCAGCTCTCCACT
              CTGCTGTTAGAGTGGCAGCTCCGAGCTGGTTGTGGCACAGTAGCTGGGGAGACCTCAGCA
              GGGCTGCTCAGTGCCTGCCTCTGACAAAATTAAAGCATTGATGGCCTGTGGACCTGCTAC
              AGTGGCCTGGTGCCTCATACTCCTCAGGTGCAGGGGCAGGGACAAGAGAAGGGGGAAGTA

22245         TCATCCCAGCAGCCTCGCACACTCCCTAGGCTTCTACCCTCCCTCCTGATGTCCCTGGAA
              CAGGAACTCGCCTGACCCTGCTGCCACCTCCTGTGCACTTTGAGCAATGCCCCCTGGGAT
              CACCCCAGCCACAAGCCCTTCGAGGGCCCTATACCATGGCCCACCTTGGAGCAGAGAGCC
              AAGCATCTTCCCTGGGAAGTCTTTCTGGCCAAGTCTGGCCAGCCTGGCCCTGCAGGTCTC
              CCATGAAGGCCACCCCATGGTCTGATGGGCATGAAGCATCTCAGACTCCTTGGCAAAAAA
              [C,-]
```

FIGURE 3M

```
        GGAGTCCGCAGGCCGCAGGTGTTGTGAAGACCACTCGTTCTGTGGTTGGGGTCCTGCAAG
        AAGGCCTCCTCAGCCCGGGGGCTATGGCCCTGACCCCAGCTCTCCACTCTGCTGTTAGAG
        TGGCAGCTCCGAGCTGGTTGTGGCACAGTAGCTGGGGAGACCTCAGCAGGGCTGCTCAGT
        GCCTGCCTCTGACAAAATTAAAGCATTGATGGCCTGTGGACCTGCTACAGTGGCCTGGTG
        CCTCATACTCCTCAGGTGCAGGGGCAGGGACAAGAGAAGGGGGAAGTAACCCCATCAGGG

22375   ACAAGCCCTTCGAGGGCCCTATACCATGGCCCACCTTGGAGCAGAGAGCCAAGCATCTTC
        CCTGGGAAGTCTTTCTGGCCAAGTCTGGCCAGCCTGGCCCTGCAGGTCTCCCATGAAGGC
        CACCCCATGGTCTGATGGGCATGAAGCATCTCAGACTCCTTGGCAAAAAACGGAGTCCGC
        AGGCCGCAGGTGTTGTGAAGACCACTCGTTCTGTGGTTGGGGTCCTGCAAGAAGGCCTCC
        TCAGCCCGGGGGCTATGGCCCTGACCCCAGCTCTCCACTCTGCTGTTAGAGTGGCAGCTC
        [C,T]
        GAGCTGGTTGTGGCACAGTAGCTGGGGAGACCTCAGCAGGGCTGCTCAGTGCCTGCCTCT
        GACAAAATTAAAGCATTGATGGCCTGTGGACCTGCTACAGTGGCCTGGTGCCTCATACTC
        CTCAGGTGCAGGGGCAGGGACAAGAGAAGGGGGAAGTAACCCCATCAGGGAGGAGTGGAG
        GGTGCCTGAGCCGCCATGTGGGCATTGGGGGAGTGATGGGAATGCCAGCAGTGATGACGT
        TGACTACTGACTGAGCACCCACTACTATGACTGAGCACTCACTCGCTAGATACTATCTTG

23042   GCCGGGCGTGGTGGCTCACGCCTGTAATCCCAGCACTTTAGGAGGCCAAGGCAGGTGGAT
        CACAAGGTCAGGAGTTTGAGATCAGCCTGGCCAACATGGTGAAACTCCATCTTTACTAAA
        AATACAAAAAATTAGCCAGGCATGGTGTTGCATGCCTGCATGCCTGTAATCCCAGTTACT
        TGGGAAGCTGAGGCAGGAGAATTGCTTGAACCCTGGAGGCGGAGGTTGTAGTGAGCCGAG
        ATCACGCCATTGCACTCCAGCCTTGGGCAAGAAGAGAAACACTCTCAAAAAAAAAAAAAAA
        [A,-,T]
        CAGGAAACTGGTGCTCAAAAAGGAAAAGTGACTCACCAAGGTCACAGACTAGGCAGTGAT
        GCTGGGGGAACCTGGCTCAGGGGACACAGACCTGGCCTGGGGCAGCCTTGCAGCTCCTCC
        ACTAAAATACTGAAAATGAGGGGCTTCGATGATGGTTATAATCGTATGGCAGAGCCCCAA
        CTCAACTGGAGCCCTGGGACCCAGAAGCTAGGGTCTCACTCCCTGCTTTTCCACAAGGCA
        CCATTAGGGCATCACCCCAGGCCTCGGCAGCCACGACGCAGGGATCCTGCCTCTCATTGG

23344   AGGAAACTGGTGCTCAAAAAGGAAAAGTGACTCACCAAGGTCACAGACTAGGCAGTGATG
        CTGGGGGAACCTGGCTCAGGGGACACAGACCTGGCCTGGGGCAGCCTTGCAGCTCCTCCA
        CTAAAATACTGAAAATGAGGGGCTTCGATGATGGTTATAATCGTATGGCAGAGCCCCAAC
        TCAACTGGAGCCCTGGGACCCAGAAGCTAGGGTCTCACTCCCTGCTTTTCCACAAGGCAC
        CATTAGGGCATCACCCCAGGCCTCGGCAGCCACGACGCAGGGATCCTGCCTCTCATTGGT
        [T,C]
        GGGGGCTTAGGGGCTCTGGGCTGCCCTCTTGAAGAGGGGGGTTCAGCCCAGCGAGGCACCC
        CCTATGCTGCACCCCACCAAGGTTAGGAAGAGGTCCTGTCCTCAGTGGGGCCCTCTGATG
        AACAGCCCATCAGGTCTGCGTCCACATGCCTTGGAAGAGATGGTGACATACTCAAAGTCC
        TTGAAGCCGCATATTAAACCACCTAGAGCACCATCTTCAAACATTTAGGGTCTGAGAAGA
        TAGGGGAAGTAAGCAATTTAAAACATTTCTTTATATTGGGCCAGGTGCAATGGCTCACGT

23873   GGTCTGAGAAGATAGGGGAAGTAAGCAATTTAAAACATTTCTTTATATTGGGCCAGGTGC
        AATGGCTCACGTCTGTAATCCCAGCGCTTTGGGAGGACGAGGATCACCTGAGGTCAGGAG
        TTCAAGATCAGCCTGGCCAACATGGAGAAACCCCATCTCTACTAAAAATACAAAAATTAG
        CTCAGGCGTGGTGATGTGCACCTGTAATCCTAGCTATTCAGGAGGCTGAGGCACAAGAAT
        TGCTTGAGTCAATATTGCACCACTGCACTCCAGCCTGGGCAACAGCGAGACTCTTGTCTC
        [A,-]
        AAAAAAAAAAAAGATATTTGCTGAAAAGACCCAGCCTGCCAAACTCAGGGGGCAGCCAAGG
        GAGGTAGTGAAATGGAAGTTGGAGCTCAGCGCTCCCACACCTCCACTGCCCTCAGGCCTT
        CTCTGCCTCTTTCCCATCAGTCAGCTGCTTCTGGGCATGGTCCTGGCAGAGACTTGGCCT
        CCTTCCAGTTCAAGCTCCCTCTTAGATTGTGTCCCACGCCACTGAGTCTTTGGGACACTG
        GGTCAGATGTCTAGTCTGGCACAATTGGCAGGAATCCCAAGAAACAGTGTGAGTGAGGGG

24764   ATAGCTGGGATTACAGGTGTGCACCACCATGCCCAGCCTAATTTTTGTATTATTAGTAGA
        GATGGGGTTTCACCATGTTGTCCAGGCTGGTCATGAACTCCTGACCTCAAGTGATCCACC
        CGCTTTGGCCTCCCAAAGTGCTGGGATTACAAGCATGAGCCACAGTGCCTGGCCTGACCC
        TGCTCTTTTGAAAGACCATTCCCCCAAATTCTGTGCACCTGTGTGCCTTTCTTCTCTCTG
        CCTCCTCTCAGCTCTGCCCCGCTCTCCTCCCTTCTCCTCTGGCAAATCCCACTCATCTCT
        [G,T]
        GAAGCCCTTCTTCCAGGGGAAGCCCTGATCATGCTGCTTTCTCCTGTGGGAGGGATGAAG
        GACGTGGCCCACGGAGTTTGTTTTGTTTTGTTTTGAGATGGAGTTTTGCTCATGTTGCCC
        AGGCTGGGGTACAATGGTACGATCTCAGCTCACTGCAACCTCTACGTCCCGGGTTCAAGC
        GGTTCTCCTGCCTTAGCCTCCCCAGTAGCTGGGATTACTGGCATGAACCACCACACCTGG
        CTAATTTTGTGTTTTTAGTAGAGATGGGGTTTCTTCATGTTGGTCAGGCTGGTCTCGAAC

24939   GACCCTGCTCTTTTGAAAGACCATTCCCCCAAATTCTGTGCACCTGTGTGCCTTTCTTCT
        CTCTGCCTCCTCTCAGCTCTGCCCCGCTCTCCTCCCTTCTCCTCTGGCAAATCCCACTCA
        TCTCTTGAAGCCCTTCTTCCAGGGGAAGCCCTGATCATGCTGCTTTCTCCTGTGGGAGGG
```

FIGURE 3N

```
        ATGAAGGACGTGGCCCACGGAGTTTGTTTTGTTTTGTTTTGAGATGGAGTTTTGCTCATG
        TTGCCCAGGCTGGGGTACAATGGTACGATCTCAGCTCACTGCAACCTCTACGTCCCGGGT
        [T,C]
        CAAGCGGTTCTCCTGCCTTAGCCTCCCCAGTAGCTGGGATTACTGGCATGAACCACCACA
        CCTGGCTAATTTTGTGTTTTTAGTAGAGATGGGGTTTCTTCATGTTGGTCAGGCTGGTCT
        CGAACTCCCAACCTCAGGTGATCTGCCTGCCTCGGCCTCCCAAAGTACTGGGATTACAGG
        GTTGAGCCACTGTGCCTGGCCCAGGCCCACGGAGTTTTAAGAGGCTTCCTGTGGCAGTGG
        CATCCAGACGGAGTGCAGAAACTCAAAGTTGAAGGCCAGAAGCTCAGGGAAGGGGGAGTG

24945   GCTCTTTTGAAAGACCATTCCCCCAAATTCTGTGCACCTGTGTGCCTTTCTTCTCTCTGC
        CTCCTCTCAGCTCTGCCCCGCTCTCCTCCCTTCTCCTCTGGCAAATCCCACTCATCTCTT
        GAAGCCCTTCTTCCAGGGGAAGCCCTGATCATGCTGCTTTCTCCTGTGGGAGGGATGAAG
        GACGTGGCCCACGGAGTTTGTTTTGTTTTGTTTTGAGATGGAGTTTTGCTCATGTTGCCC
        AGGCTGGGGTACAATGGTACGATCTCAGCTCACTGCAACCTCTACGTCCCGGGTTCAAGC
        [G,A]
        GTTCTCCTGCCTTAGCCTCCCCAGTAGCTGGGATTACTGGCATGAACCACCACACCTGGC
        TAATTTTGTGTTTTTAGTAGAGATGGGGTTTCTTCATGTTGGTCAGGCTGGTCTCGAACT
        CCCAACCTCAGGTGATCTGCCTGCCTCGGCCTCCCAAAGTACTGGGATTACAGGGTTGAG
        CCACTGTGCCTGGCCCAGGCCCACGGAGTTTTAAGAGGCTTCCTGTGGCAGTGGCATCCA
        GACGGAGTGCAGAAACTCAAAGTTGAAGGCCAGAAGCTCAGGGAAGGGGGAGTGTGAGTT

25092   ATCATGCTGCTTTCTCCTGTGGGAGGGATGAAGGACGTGGCCCACGGAGTTTGTTTTGTT
        TTGTTTTGAGATGGAGTTTTGCTCATGTTGCCCAGGCTGGGGTACAATGGTACGATCTCA
        GCTCACTGCAACCTCTACGTCCCGGGTTCAAGCGGTTCTCCTGCCTTAGCCTCCCCAGTA
        GCTGGGATTACTGGCATGAACCACCACACCTGGCTAATTTTGTGTTTTTAGTAGAGATGG
        GGTTTCTTCATGTTGGTCAGGCTGGTCTCGAACTCCCAACCTCAGGTGATCTGCCTGCCT
        [C,T]
        GGCCTCCCAAAGTACTGGGATTACAGGGTTGAGCCACTGTGCCTGGCCCAGGCCCACGGA
        GTTTTAAGAGGCTTCCTGTGGCAGTGGCATCCAGACGGAGTGCAGAAACTCAAAGTTGAA
        GGCCAGAAGCTCAGGGAAGGGGGAGTGTGAGTTGAGGAGTCTCTTGGCTGCCAGGGCCAG
        AAACCGAACTCCAAGCCTCTCCACAACAGCGGGTGTAGAGCATGTAGAATCAGAGAGGAG
        GCTGAGCCATGCAGCCCCGAGAAGAGGGGAATGCCACTGAGCCACAGAGACCCAGTGCCA

25428   AGTGCAGAAACTCAAAGTTGAAGGCCAGAAGCTCAGGGAAGGGGGAGTGTGAGTTGAGGA
        GTCTCTTGGCTGCCAGGGCCAGAAACCGAACTCCAAGCCTCTCCACAACAGCGGGTGTAG
        AGCATGTAGAATCAGAGAGGAGGCTGAGCCATGCAGCCCCGAGAAGAGGGGAATGCCACT
        GAGCCACAGAGACCCAGTGCCACTGCCAGGTGTCTCTGCCTCCACTTCCCATGACCC
        [T,G]
        GCCTGTCTCTGTATGCAGGCTTCACCCTCTCTCGTTGTACATTGTACACATTCTAGGTGA
        CACCAGCAGCTTCTGATTCTCATCTCCCATAACATCAGCCCCCCAGAGAGGGGACAACTG
        CTGAGCTGATAACATAATAGATGCCCCTTTCCTGGAGGCCATGGTCATGGTCAGCGTGGA
        GAGGATGAAGCCTGAGCAGGCAGGATCGGGGGTCTAGAGGGGAAGGAGGTGGAAGTT

25513   GGCCAGAAGCTCAGGGAAGGGGGAGTGTGAGTTGAGGAGTCTCTTGGCTGCCAGGGCCAG
        AAACCGAACTCCAAGCCTCTCCACAACAGCGGGTGTAGAGCATGTAGAATCAGAGAGGAG
        GCTGAGCCATGCAGCCCCGAGAAGAGGGGAATGCCACTGAGCCACAGAGACCCAGTGCCA
        CTGCCAGGTGTCTCTGCCTCCACTTCCCATGACCCGGCCTGTCTCTGTATGCAGGCTTCA
        CCCTCTCTCGTTGTACATTGTACACATTCTAGGTGACACCAGCAGCTTCTGATTCTCATC
        [C,T]
        CCCATAACATCAGCCCCCCAGAGAGGGGACAACTGCTGAGCTGATAACATAATAGATGCC
        CCTTTCCTGGAGGCCATGGTCATGGTCAGCGTGGAGAGGATGAAGCCTGAGCAGGCAGGA
        TCGGGGGTCTAGAGGGGAAGGAGGTGGAAGTTGAGATCACAGACCTGTGGTCAGGTGGCC
        TGGGAAGGGTTTGACGAGTGTCGGCCCAAAGAGCTTGGAAGGGATTTTGCTGCTGTGGGT
        GAGCACTGCCTCTCCCCTTAGGGACAACAGCCACCTCTTCTCTCCCCATTTGCCTTTCCC

25684   CCAGTGCCACTGCCAGGTGTCTCTGCCTCCACTTCCCATGACCCGGCCTGTCTCTGTATG
        CAGGCTTCACCCTCTCTCGTTGTACATTGTACACATTCTAGGTGACACCAGCAGCTTCTG
        ATTCTCATCTCCCATAACATCAGCCCCCCAGAGAGGGGACAACTGCTGAGCTGATAACAT
        AATAGATGCCCCTTTCCTGGAGGCCATGGTCATGGTCAGCGTGGAGAGGATGAAGCCTGA
        GCAGGCAGGATCGGGGGTCTAGAGGGGAAGGAGGTGGAAGTTGAGATCACAGACCTGTGG
        [C,T]
        CAGGTGGCCTGGGAAGGGTTTGACGAGTGTCGGCCCAAAGAGCTTGGAAGGGATTTTGCT
        GCTGTGGGTGAGCACTGCCTCTCCCCTTAGGGACAACAGCCACCTCTTCTCTCCCCATTT
        GCCTTTCCCTTCTGTAGATATGAAACACAGGCCTCCTTGTCAGGCCCCTACTTAACCTCC
        GTGATGGGGAAAGCGGCCGGAGAAAGGGAGTTCTGGCAGCTTCTCCGAGACCCCAACACC
        CCACTGTTGCAAGGTGAGTCATGGCCTGACACTCTGGATGTGTCCCCTACCCCAAGCTTA

26165   GTGATGGGGAAAGCGGCCGGAGAAAGGGAGTTCTGGCAGCTTCTCCGAGACCCCAACACC
        CCACTGTTGCAAGGTGAGTCATGGCCTGACACTCTGGATGTGTCCCCTACCCCAAGCTTA
```

FIGURE 3O

CTCAGCCAAGAGGCTTCATCAACTCACCCCAGCTTTCCCTAGCACCCTCCTGGGCCACAC
CTTCACAAAATCACTGATGCTCAAAGTTGGATATAATATATTGAACTGAAGCCTTAGCAT
TTTTATGCAAGTTACTGTGGAAATTCTAGGAAACCAGACAGATTACAAAAAAAAAAAAAA
[A,-]
CTAGAAGAAAATTAACATCACCTAGGATATACTACCTAGGAATAACGTCTTTTATTTTGA
GATGGAGTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAGCGGTATGATCTCGGCTCGCTGC
AACCTCCGCCTCCTGGGTTCATGTGATTCTTCCACCTCGGCCTTCCTAGAGCCCAAGTGG
TCTGCCTGCCTCTGCCTCCCAAAGTTCTGGGATTACAGGCATGAGCCACCGCACCCAGCC
AAAATTACTTAACTTTTCTTCTAGATACTTTTTAAAAATATGGCAGTAAGTTTTTCATAA

FIGURE 3P

FEATURES:
Exon: 2132-2195
Intron: 2196-5670
Exon: 5671-5863
Intron: 5864-7672
Exon: 7673-7761
Intron: 7762-9149
Exon: 9150-9302
Intron: 9303-10101
Exon: 10102-10274
Intron: 10275-10399
Exon: 10400-10586
Intron: 10587-12128
Exon: 12129-12294
Intron: 12295-25922
Exon: 25923-26040

Allelic Variants (SNPs):

DNA Position	Major	Minor	Domain	Protein Position	Major	Minor
1022	C	T	Beyond ORF(5')			
1882	C	T	Beyond ORF(5')			
1951	G	A	Beyond ORF(5')			
2940	A	G	Intron			
3831	G	A	Intron			
6732	G	A	Intron			
7558	G	A	Intron			
7931	T	A	Intron			
8190	T	C	Intron			
8652	C	T	Intron			
9370	G	C	Intron			
9463	C	T	Intron			
9883	G	T	Intron			
11594	G	A T	Intron			
12361	A	G	Intron			
13187	T	C	Intron			
13284	T	C	Intron			
21341	C	G	Intron			
21796	A	G	Intron			
22159	C	T	Intron			
22334	C	A	Intron			
22346	C	–	Intron			
22476	C	T	Intron			
23143	A	– T	Intron			
23445	T	C	Intron			
23974	A	–	Intron			
24865	G	T	Intron			
25040	T	C	Intron			
25046	G	A	Intron			
25193	C	T	Intron			
25529	T	G	Intron			
25614	C	T	Intron			
25785	C	T	Intron			
26266	A	–	Beyond ORF(3')			

Context:
DNA
Position

1022   TTGGAGATATTTTAAGGTCATAGTGTCTTCACAAATTGAGCTGAAAGGGAACTGTTAGGA
        TGATCTTGCCTAACCCTCTCATCTCACACAGGAAGAACTATTTTAAACTCGAGAGGTTAA
        GTGACCTGGCCAAAGTCACACAGCCACCACTAGTTAACTCGTATACATTGATTCTCCTGT
        GGGGCTGGGCAGATGAGGAATCTTTTGTTCTCTTCCCTGTTTGCAGAGATTTTTTTTGAG
        GTTACTTTCCGAGTTCTGGCAAGTACCCCTGCTTCTGGTAGCTTTGTGTCTCGATTCAAT

FIGURE 3Q

```
                [C,T]
                TCATTCTTTTTATTTTATTTTATTTTTGAGACAGGGTCTCACTTTGTCACCCAAGCTGGA
                GTGCAGTGGTGTAATCTTGGCTCACTGTAGCCTCCACCTCTTGGGTTCAAGCGATCCTCC
                TGCCTCAGCCCCCCAAGTAGCTGGGATTACAGACGTCTGCCACCACGCCAGGCTAATTTA
                TGGTTTTTTGTATGTGTTTTTTGTGTTTTTGTAGAGACAGTGTTTCCCCATGTTGCCCAG
                GCTGGTCTCCAACTCCTGAGCTCAAGTGATCTGCCCGCCTCAGCCTTTCAAAGTGCTAGG

     1882       ACAGCTGACTCCAGCAATGCTGCTCACGTGACCACTGCAGCTGCAGCTCCCGTTCCACTC
                CTTGTCCTGGGCTAGGTGGGCACTACCAGGGGCTCCTTTGGTAAGGAGTACCGGGTAGGC
                ACCCGGTCCTGCCAATCCACCACTGGAACAGCTGGGGGGACAGCAGACAGGCACGGTCGG
                ACAGACTTGACAGATCAGGCATCAGGCCCTCTGCGCTGGTCCCGGGCTCTTTAAGCAGGA
                ACGTGAATGGCCTCAAGATGTCTCACATGGTCCCACTAGCCCTCCTCCTCCCTTTGTTCC
                [C,T]
                TACCTCCAGGAGGGCTGCTCTGCCCTTCCTTCCTCTGTTCTTTGGCCTTATGTTCCCCGC
                CACCACAGGCCTTCCCCCGCCCCACCCCTCTGCAGACTTAGCCGTGCATTGCAGGCATGG
                AGGATTAATCAGTGACAGGAAGCTGCGTCTCTCGGAGCGGTGACCAGCTGTGGTCAGGAG
                AGCCTCAGCAGGGCCAGCCCCAGGAGTCTTTCCCGATTCTTGCTCACTGCTCACCCACCT
                GCTGCTGCCATGAGGCACCTTGGGGCCTTCCTCTTCCTTCTGGGGGTCCTGGGGGCCCTC

     1951       GGCTAGGTGGGCACTACCAGGGGCTCCTTTGGTAAGGAGTACCGGGTAGGCACCCGGTCC
                TGCCAATCCACCACTGGAACAGCTGGGGGGACAGCAGACAGGCACGGTCGGACAGACTTG
                ACAGATCAGGCATCAGGCCCTCTGCGCTGGTCCCGGGCTCTTTAAGCAGGAACGTGAATG
                GCCTCAAGATGTCTCACATGGTCCCACTAGCCCTCCTCCTCCCTTTGTTCCCTACCTCCA
                GGAGGGCTGCTCTGCCCTTCCTTCCTCTGTTCTTTGGCCTTATGTTCCCCGCCACCACAG
                [G,A]
                CCTTCCCCCGCCCCACCCCTCTGCAGACTTAGCCGTGCATTGCAGGCATGGAGGATTAAT
                CAGTGACAGGAAGCTGCGTCTCTCGGAGCGGTGACCAGCTGTGGTCAGGAGAGCCTCAGC
                AGGGCCAGCCCCAGGAGTCTTTCCCGATTCTTGCTCACTGCTCACCCACCTGCTGCTGCC
                ATGAGGCACCTTGGGGCCTTCCTCTTCCTTCTGGGGGTCCTGGGGGCCCTCACTGAGATG
                TGTGGTGAGTAACTCGCCTCTATCCTGTGCCTCTTTCCTCCTGGGTCCTTAGTGGGGTGG

     2940       AACATAGGGAGACCCCATCTCTACAAAAAATAAAAAAAATTAAAAAATAGCTGGGCATGG
                GGAAGACTTTCTGAAGACCAAGAGGACACATGGGAGCTGAAACTCGAAGGAAGAAAAGGA
                GCTGGCAGGAAAGGAGTGGGGGACACACATTCTAGGCAGCAGGAAGTGAGCCTTCGGAGG
                TCCTGCCTGCTCCAGCTCTGTGCCCCAAGGGGTCTCTTGGAGCACAGTCTCCTGGGACCT
                GTCTATGAGTCTGAGCTTAGAGGCTCAGGGCTGCTCCTTCAGACAGGAGGCAGAAGGCAG
                [A,G]
                CTTTGGGAACTTTGGGCCGCCCACGCGCCTTTTCTCCTCCTCTGCACCTAGGATTACGTT
                GAGCAATACACTTTCACCCCCATGGTCTCTTGAGACCCTGGGGAAACCCTGAGAGGTGGG
                TGCAGTCATGTCCAGGTGTCAAGTGAAGAAGTCGAGGGTTGGAGGGGCTGAGTGACCCAC
                TCAGGGTGCTCCACCTTTTCCAGAGCTTTGCTGAACTTAGTTTTTAGAACTTGAAGCCTC
                GTTTGTTTTCGTTTTGTTTTTTGTTGAGAGAGGTTCTCCCTCTGTTGCCCAGGCTGGAGT

     3831       GACACCTCAGGTCTGGGCCCAGGAACCCCAGCTCTTGGTTCATGTCCGGACAGTCCCCAG
                GGGAGTTCTGGGTTCAACCAGCAAGAGCTCTTCCTCCTGGCTGATCTGGTCCTCAGCCTT
                GGACAGTTAGTCCATTAACCTGACCCCACAGGAGCCCCAATCCCTTGGGGTCTGGGGAAT
                CTTGAACTGGGGTTTGGGGTGCAAATATCTGCACTGAGTCACTTAATTGCACCCAGCCTC
                ATTCCTTTATCTGTAAAGTGGGCTAAGAATGCTCCCCTGCCTTCCTCCTCGGTGTAGTAC
                [G,A]
                AGGAAGGATCCCATGACACCTGCTCTCCCAGTTTAAAGCTCTATATGTATGTTGTGAAAT
                TGACAGGGATCGCTGCACAAACGCTAATGCAAAGTGGCTCCTGTGCTTCCTTTTCTCTT
                TCTTCTTCTTTTTTTTTTTTTTTTAATTTTCTTCTAGAGATGAGGTCTCACTATATTGCCCA
                GGGTTGGTTTCAAACTCCTAGGGTCAAGCGATCCTCCCACCTTGGCCTCCCAAACTGCTG
                GTATTACAGGCGTGAGCCACTCTGTCTGGCTCCTATGCTTGTGAATGTCAACAGCAATCA

     6732       TGAATAGCTGGGATTACAGGCGTGTGCCACCATGCCCAGCTGATTTTTGTATTTTTGGTA
                GAGATAGGGTTTCACCATGTTGGCCAGGCTGGTCTTGAACTCCTGACCTCAGGTGATCCG
                CCTGTCTTGGCTTCCCAAAGTGTTGGGATTATAGGCATGAGCCACTGCACCAATCCAAAA
                GCAGCATCTTTGTGCTCCCTTTTCAAGAGGCATCACAGAGAGGCCTGTTTTGGGGTTTGA
                ATGAGAGGCGAAGAATCAGCCATGGAGTGCCTCTTTCTCAGACTCCCTCTTGAGAAGTGG
                [G,A]
                TGCAGGGGTGGAGAGAAAAGAAGACTAGGCATAGTGGCTCATACCTGTAATCCCAACATT
                TTGGGAGGCTGAGGCAGGAAGATTGCTTGAGCTCAGGAGTTTGAGACCAGCCTAGGCAAC
                ATAGTGAGACCACATCTCTTAAAAAAAAAGAAAAAGAAAAAAAAATGAGCCAGGTGTAGTGA
                CTCATGCCTGTGGTCCCCACTTCTCCGGAGGCAAAGGTGGGAGGATCTTTTGAGGCTGAG
                AAATCGAGGCTACAGTGAGCCATGGTGGCACCACTGCACTCCAGCCTGGGAGACAGAGAG

     7558       AAGAATCGCTTGAACTCGGGAGGCGGAGGTTGCAGTGAGCTGAGAACATGCCACTGCACT
                CCAGCCTGGGCAACAAGAGCGAAACTCTGTCTCAAAGAAAATAAATAAATAAAATAAAAA
```

FIGURE 3R

```
          AATAAAAAAGGAGGGGGCATATGGGTGAAGTATGGACAAAATAGTGGGGCAGGCACAGAT
          GATCTGGACACAGGAGCCCTTGGAGTTTATTCTTGAATCTAACTGTTCATCTTTATTAAA
          TATTTGTGGCATACACCTCACAACAACATAGCCAACACACCTCCTTTTGGAGCTTTTATC
          [G,A]
          AAGTTTCCCACTGTTAAGATTTTTTCCCGCTTTGTGATGCGGGTGGGGTGGGTGCTGTAA
          GCAGGCTTACGGGGTGGCAGTTTCTCACAAAGGCATTAACTGGCCTTGTCCTAGGTCTGC
          CTTCAGCGAGGATGACGGTGACTGCCAGGGCAAGCCTTCCATGGGCCAGCTGGCCCTCTA
          CCTGCTCGCTCTCAGAGCCAACTGTGAGTTTGTCAGGGGCCACAAGGGGGACAGGCTGGT
          CTCACAGCTCAAATGGTTCCTGGAGGATGAGAAGAGAGCCATTGGTGAGCAGACACCATC

7931      GGTGGCAGTTTCTCACAAAGGCATTAACTGGCCTTGTCCTAGGTCTGCCTTCAGCGAGGA
          TGACGGTGACTGCCAGGGCAAGCCTTCCATGGGCCAGCTGGCCCTCTACCTGCTCGCTCT
          CAGAGCCAACTGTGAGTTTGTCAGGGGCCACAAGGGGGACAGGCTGGTCTCACAGCTCAA
          ATGGTTCCTGGAGGATGAGAAGAGAGCCATTGGTGAGCAGACACCATCCGCTGGGGGTGG
          GGAGCAGCTGGGAGGGCTCATCAGATGATATTCTCCAATGAGAATCAGAACTTTGGGTTT
          [T,A]
          CTCCCCAGGCGTCTTTCCCACCATCCATTCTGCCCATCTCACTGCCTACGTAGAGGCTCG
          AACCTGTCCCCATAGCCATCCTTGACCCAGCTTTTCCCGCGCTGCACACATACTATTGAC
          AGGTGTGTTTCGTGGTTTTTTGTTTTTTGTTTGTTTGTTTGTTTTGAGTTGGAGGTTTGC
          TCTTGCTGCCCAGGCTGGAGTACAATGGCGCAATCTCAGCTCACCGCAATCTCTGCCTCC
          TGGGTTCAAGCAATTCTCTTGCCTCAGCCTCCTGAGTAGCTGGGATTACAGGCATGCGCC

8190      ATCAGATGATATTCTCCAATGAGAATCAGAACTTTGGGTTTTCTCCCCAGGCGTCTTTCC
          CACCATCCATTCTGCCCATCTCACTGCCTACGTAGAGGCTCGAACCTGTCCCCATAGCCA
          TCCTTGACCCAGCTTTTCCCGCGCTGCACACATACTATTGACAGGTGTGTTTCGTGGTTT
          TTTGTTTTTTGTTTGTTTGTTTGTTTTGAGTTGGAGGTTTGCTCTTGCTGCCCAGGCTGG
          AGTACAATGGCGCAATCTCAGCTCACCGCAATCTCTGCCTCCTGGGTTCAAGCAATTCTC
          [T,C]
          TGCCTCAGCCTCCTGAGTAGCTGGGATTACAGGCATGCGCCACCACACCCAGCTAATTTT
          GTATTTTTAGTAGACGTGGGGTTTCTCCATGTTGGTCAGGCTGGTCTCGAACTCCTGACC
          TCAGGTGATCCGCTTGCCTTAGCCTCCGAAAGTGCTGGGATTACAGGCATGAGCCACTGC
          GTTAGGCCCACTGACAAGCCTTGTATTGGCTAGCCACCAAGATTGACTTGATTATCCACC
          TTCGGGACAACTGGACAGCCTGCTTATGACTTACGCCATAGTCTGTCTCTACTAGCTCTC

8652      TACAGGCATGAGCCACTGCGTTAGGCCCACTGACAAGCCTTGTATTGGCTAGCCACCAAG
          ATTGACTTGATTATCCACCTTCGGGACAACTGGACAGCCTGCTTATGACTTACGCCATAG
          TCTGTCTCTACTAGCTCTCCTGCCCTGACTTGACCCAGCATACAACAGCCAGAGCCAGCC
          TTTTCAATATAAACCTGATCTTGCTGGCACTGCTTAAACCCTGCAGGGGCCCTCGCACTGC
          TCCATGGCCCAGCCTGTCTACCCTTACCTTCTGCCCAGGCTCTGCTCATCCATTCTCTGC
          [C,T]
          TCCCACACACCTGCCCTCTGTGGGCTCCAGCCATACCATCTCTCAACTCATAAGCCAGTT
          TTTTCATACAGGCTCCCTCCATCTGGACTGGCTTCCCTGCGTGCAGTTCACTCCTGCTCT
          ACCTTTGGCTCTGCCTCCACCCATCCTCAGCCGTCTCCAGCATTACCTCCTTGGAGAATC
          CTGCCTTGACTTCCCAGCCACCCAAATATCACTACTTGGTCTGCATTCTCGTTGCAATTG
          CAGTCGCATGAGCAATTGCTGTGGTTGAGGCCCGAACTGCGCAAGTGCCTGTCTGCCATG

9370      AGGCCAGGGTCCCAGGTGCTGGCGGGGCTGGCTGCTGGGTGGGGGCAGAGAGGCAACCCC
          TCTGTTTTTTTCCCTCTCAGGGCATGATCACAAGGGCCACCCCCACACTAGCTACTACCA
          GTATGGCCTGGGCATTCTGGCCCTGTGTCTCCACCAGAAGCGGGTCCATGACAGCGTGGT
          GGACAAACTTCTGTATGCTGTGGAACCTTTCCACCAGGGCCACCATTCTGTGGGTGAGTA
          GGTCAGACCGTGCCAAGGCCAGGCTGGCACTCCCTCAGTCCCCAGGTCTGCACTGATGAC
          [G,C]
          TCCATACCCTGGCCCCCACACTCACCTTTCCTTGGGGCTCCTCCGAATCAAGTCCTTTAG
          GGACGAATTGGCGAGGGCTCATGGGTGATGCTCCAGCTGTGAGCCAGCTTTGGAGCTGGT
          AGGTGGATCTCTTGAGGCCAGGAGTTCAAGACAACGTGGTGAAACCCCATCTCTACTAAA
          AATAAAAAAGTTAGCCGGGCATGGTGGCACATGCCTGTAGTCCCAGCTACTCGGGAGGCT
          GAGGCAGGAGAATCACTTGAACCTGGGAGGCGGAGGCTGCAGTGAGTGGAGATCGCACCA

9463      GGGCCACCCCCACACTAGCTACTACCAGTATGGCCTGGGCATTCTGGCCCTGTGTCTCCA
          CCAGAAGCGGGTCCATGACAGCGTGGTGGACAAACTTCTGTATGCTGTGGAACCTTTCCA
          CCAGGGCCACCATTCTGTGGGTGAGTAGGTCAGACCGTGCCAAGGCCAGGCTGGCACTCC
          CTCAGTCCCCAGGTCTGCACTGATGACGTCCATACCCTGGCCCCCACACTCACCTTTCCT
          TGGGGCTCCTCCGAATCAAGTCCTTTAGGGACGAATTGGCGAGGGCTCATGGGTGATGCT
          [C,T]
          CAGCTGTGAGCCAGCTTTGGAGCTGGTAGGTGGATCTCTTGAGGCCAGGAGTTCAAGACA
          ACGTGGTGAAACCCCATCTCTACTAAAAATAAAAAAGTTAGCCGGGCATGGTGGCACATG
          CCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATCACTTGAACCTGGGAGGCGG
          AGGCTGCAGTGAGTGGAGATCGCACCACTGCCCTCCAGCCTGGGCAACAGAGTGAGTGAG
          ACTCTGTCTCAAAAAATAAAAAATAAAATAAAACTCCCTAGTGATTCCAATGTGCAGCT
```

FIGURE 3S

```
9883      GCCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATCACTTGAACCTGGGAGGCG
          GAGGCTGCAGTGAGTGGAGATCGCACCACTGCCCTCCAGCCTGGGCAACAGAGTGAGTGA
          GACTCTGTCTCAAAAAATAAAAAATAAAATAAAACTCCCCTAGTGATTCCAATGTGCAGC
          TAAGTTTGGAAATAGGTGGTATGGGGTCAAGTCCTCTTGGGCCTCCCTCCTCCAGTCCTT
          CTCCCTAACCTCTAGCCCTCAAGTTGCAGAGTGATCAGCCAAACCAGTTTGCCCAGAAAT
          [G,T]
          AGCAGTTTCCTGGGACACAGGATTTTCAGAGTCCAGACAAGGAAAGTCTTGGGCAGACCA
          GGTTGAGTTGGTGCCCTTAGCTGATCTGACCATGTTGCCCTTCTTCTCCAAGCCCTCCTG
          TGGTTGTCCATAGCTACAAGGGCCTGACCCTCAAGCCCCTGCCTGTCCTGGCCCCTTTGG
          CTCTCCAGCTCATTGCATGTTCTGTCCCCCACTTCAAGACACAGCAGCCATGGCAGGCTT
          GGCATTCACCTGTCTGAAGCGCTCAAACTTCAACCCTGGTCGGAGACAACGGATCACCAT

11594     AAAAAAAAAAAATGGAGAAGAAGGAAGCTGGACATGGTGGCTCGTGCTTATAATCCTAGCA
          CTCTGGGAAGCTGAGGCAGATGGGATTGCCTGAGCCCAGGAGTTTGAGACCAGCCTGGGCA
          ACATGGTGAAACCCTGTCTTTACTAAAATACGAAAGATTAGCCAGGCATGGTGGTAGACA
          CCTATAATCCCAGCTACTAGGGAGGCTGAGCCACAAGAATCACTTGAACCTGGGAGACAG
          AGGTTGCAGTGAGCCGAGATCGCGCCATTGCACTCCAGCCTGGGCGACAGTGTGAGACTC
          [G,A,T]
          GTCTCCAGAAAAAACAAGAATGGATAGAGTGGAGCCAAGAAGAGGCAGGAAGAACAAAGA
          CACAGAGGTGCACAGAGTTTGGGGGAATTTTGAGGAATGGTCTTGCAAAAGAGTGGGATC
          TGGGAGAATGAGTGGGAGTGGAAAGCAGATGAATGAAGGAAGGTGAGCGCATCAGGGTA
          ACAGAGATGCGTTGTGAACAAATGCATGTTCTAGGAAGAGCCCTCTGGAGTGCTAGGTGC
          CAGAGAGGTGGGAGGAAGGATACTGGAAGCAGAGAAACCAGTGAGGGGCCTGATCTTGGG

12361     ACAAGAAGACAAATAATCCAGGCTCTCTGTCCTCACACCAGCTGCCCGCCCCTTTCTTCC
          TGGCACAGTCATGTTGGAACCAGCTGCTGAGACCATTCCTCAGACCCAAGAGATCATCAG
          TGTCACGCTGCAGGTGCTTAGTCTCTTGCCGCCGTACAGACAGTCCATCTCTGTTCTGGC
          CGGGTCCACCGTGGAAGATGTCCTGAAGAAGGCCCATGAGTTAGGAGGATTCACGTGAGA
          CTCCCACCTCCCAGTCCTCACCCCACCCAACCTCACATGCCTGATAACAGGGTCACAGAA
          [A,G]
          AGACGGGGAACAGAGGAGAGGGTTCCCTCGGGAGAGACACTGGCCCTGCTTCTGCTTCTA
          CCTGCTCAGCTCCTTTCTTGCCCACGGTGTTATGGAAACAGGGAGCCATAGGCCAGCATT
          GTCACTGAGAGAGCAGGCTTTGGAGGCAGAGCCCCCCAGTTGGAATCCCAACTCTAACCA
          GCTAGGTTCCAGGTAGGCACCCACAATTCACCGAGGAGAACAGTTGTGCCCCTTCCCTGC
          AGGGCCAGTGTGAAGAGTCCAGGAGTTAGTACACATAGAGATAGTGGCATGTGCTTTTTA

13187     GGCACGTGCCACACAACCCTGGGTAATTTTTTTTTTTTTTTTTTTGAGATAGGGTCTCTG
          TCTGTTGCCCAGGCTGGTCTCAAATTCCTGGCCTCAAACCATCCTCACACCTGAGGCGCT
          CAAAATATTGGGATTATAGGTGCGAGCCATCATGCTCAGCCAGAATAATAACTGGTTTTT
          TTTGTTTTTTTTTTGAGACAGAGTCTCACTCTATTACCCAGGCTCTGGAGGCCCAACTCG
          TGTTTGTGTATTTGTTTATTTTTATTTATTTATTTATTTCGAGACAGAGCCTCTCTCTTT
          [T,C]
          ACCTAGGCTGGAGTGCAGTGGCGCAATCTCGGCTCACTGCAACCTCCGTCTCCTGGGTTC
          AAGTGATTGTCCTGCCTCAGCCTCCTGAGTAGCTGGTGCTACAGGCGCGTGCCACCATGC
          CCAGCTAATTTTTGTATTTTTAGTAGAGACAGGGTTTTACTATGTTGGCCAGCTGGTTTC
          TAACTCCTGAACTCGGGTGATCTGCCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGC
          ATGGGCCTCCGTGCCCGGCCATGTATTTATTTAGGCAAGGTCTCTCTCTGTTATCCAGGC

13284     ACCATCCTCACACCTGAGGCGCTCAAAATATTGGGATTATAGGTGCGAGCCATCATGCTC
          AGCCAGAATAATAACTGGTTTTTTTTTTGTTTTTTTTTTTGAGACAGAGTCTCACTCTATTAC
          CCAGGCTCTGGAGGCCCAACTCGTGTTTGTGTATTTGTTTATTTTTATTTATTTATTTAT
          TTCGAGACAGAGCCTCTCTCTTTCACCTAGGCTGGAGTGCAGTGGCGCAATCTCGGCTCA
          CTGCAACCTCCGTCTCCTGGGTTCAAGTGATTGTCCTGCCTCAGCCTCCTGAGTAGCTGG
          [T,C]
          GCTACAGGCGCGTGCCACCATGCCCAGCTAATTTTTGTATTTTTAGTAGAGACAGGGTTT
          TACTATGTTGGCCAGCTGGTTTCTAACTCCTGAACTCGGGTGATCTGCCTGCCTCGGCCT
          CCCAAAGTGCTGGGATTACAGGCATGGGCCTCCGTGCCCGGCCATGTATTTATTTAGGCA
          AGGTCTCTCTCTGTTATCCAGGCTGAAGTGCAGTGGCACATTCATAGCTCACTGCAGCCT
          CAAATTATCCAAGTAACAGGGACTACAGGCATGCACCACCACACCCATCTACTTTTTTTT

21341     TCAGCTCACTGCAACCTCTGCCTCCCTGGCTTAAGCGATCCTCCCACCTCAGCCTCCTGA
          GTACGTGTGACCATAGGCCCATGGCACAAAGCCCAGCTAATTTTTTGTATTTTTAGTAGA
          AATGTGGTTTCACCATGTTGCATAGGCTGGTCTCGAACTTCTGAACTCAAGTGATCTGCC
          TGCCTTGGCCTCCCAAAGTGCTGGGATTCTAGGTATGAGCCACCCGCTCGGCCTATAAT
          GGCACTTTCCTATCCCATTGATGAGGCTCTACTCTCATGACCTAATCATCTCCCAAAGGC
          [C,G]
          CTAAGGCCTCCTGATACCATCACCTTTGGGGTTAGGTTTTAACATATACATTTTGGGGGG
          ACACAGACATTTTAGACCATAGCACCTCCATTGAAAGGAAACATTTCTGACACCTGGCTA
```

FIGURE 3T

```
            TCTCAAAGGGCCCTTTCAGTTCCCCTGCAGGCTGCATTCCCACATCACCAACAAGAGCAG
            CGACACTCACTCAGAGGTTAAATAACTTGTCCAGAGTCACAGCAGTAATGAATGACAGAG
            CTGGGGCTTGAATCCAGGCGTCCTCCTAGAGCCTGGATTCTGTGTAGTGAGTGAAAGCTG

    21796   CATTCCCACATCACCAACAAGAGCAGCGACACTCACTCAGAGGTTAAATAACTTGTCCAG
            AGTCACAGCAGTAATGAATGACAGAGCTGGGGCTTGAATCCAGGCGTCCTCCTAGAGCCT
            GGATTCTGTGTAGTGAGTGAAAGCTGACTCCTGGGAGACTTCTGCGTGGTCCTGGTTCTC
            TCTCCAGACTGCACTGCGCAAGTTTCTCTTCCTGATGGTCCCTAGGGTATTACAAAGACA
            GTGGCCCTGCCTGTCAGGTGTTTTTATTACCAGATGAGGTCATGGCCTCAGGAACCCTGT
            [A,G]
            GGAAGCTGAGTTCAGAGTCTTTGAGCAGGCTTTAGGGAGGTTCCAGCTTCCCACCACCAA
            GCCCCAGGTGGATTCTTACAGACTCTAGCCTCAGGGTGGGGGGTCTGGAAGATGAGGTTG
            CGGGGTGCGATATTCTGCCCAATTCGCCCCTCCTTGCTCAATCTGTTTCTGCAGGTATTG
            CTGACTACAGACCCAAGGATGGAGAAACCATTGAGCTGAGGCTGGTTAGCTGGTAGCCCC
            TGAGCTCCCTCATCCCAGCAGCCTCGCACACTCCCTAGGCTTCTACCCTCCCTCCTGATG

    22159   CCCAGGTGGATTCTTACAGACTCTAGCCTCAGGGTGGGGGGTCTGGAAGATGAGGTTGCG
            GGGTGCGATATTCTGCCCAATTCGCCCCTCCTTGCTCAATCTGTTTCTGCAGGTATTGCT
            GACTACAGACCCAAGGATGGAGAAACCATTGAGCTGAGGCTGGTTAGCTGGTAGCCCCTG
            AGCTCCCTCATCCCAGCAGCCTCGCACACTCCCTAGGCTTCTACCCTCCCTCCTGATGTC
            CCTGGAACAGGAACTCGCCTGACCCTGCTGCCACCTCCTGTGCACTTTGAGCAATGCCCC
            [C,T]
            TGGGATCACCCCAGCCACAAGCCCTTCGAGGGCCCTATACCATGGCCCACCTTGGAGCAG
            AGAGCCAAGCATCTTCCCTGGGAAGTCTTTCTGGCCAAGTCTGGCCAGCCTGGCCCTGCA
            GGTCTCCCATGAAGGCCACCCCATGGTCTGATGGGCATGAAGCATCTCAGACTCCTTGGC
            AAAAAACGGAGTCCGCAGGCCGCAGGTGTTGTGAAGACCACTCGTTCTGTGGTTGGGGTC
            CTGCAAGAAGGCCTCCTCAGCCCGGGGGCTATGGCCCTGACCCCAGCTCTCCACTCTGCT

    22334   CCCTGAGCTCCCTCATCCCAGCAGCCTCGCACACTCCCTAGGCTTCTACCCTCCCTCCTG
            ATGTCCCTGGAACAGGAACTCGCCTGACCCTGCTGCCACCTCCTGTGCACTTTGAGCAAT
            GCCCCCTGGGATCACCCCAGCCACAAGCCCTTCGAGGGCCCTATACCATGGCCCACCTTG
            GAGCAGAGAGCCAAGCATCTTCCCTGGGAAGTCTTTCTGGCCAAGTCTGGCCAGCCTGGC
            CCTGCAGGTCTCCCATGAAGGCCACCCCATGGTCTGATGGGCATGAAGCATCTCAGACTC
            [C,A]
            TTGGCAAAAAACGGAGTCCGCAGGCCGCAGGTGTTGTGAAGACCACTCGTTCTGTGGTTG
            GGGTCCTGCAAGAAGGCCTCCTCAGCCCGGGGGCTATGGCCCTGACCCCAGCTCTCCACT
            CTGCTGTTAGAGTGGCAGCTCCGAGCTGGTTGTGGCACAGTAGCTGGGGAGACCTCAGCA
            GGGCTGCTCAGTGCCTGCCTCTGACAAAATTAAAGCATTGATGGCCTGTGGACCTGCTAC
            AGTGGCCTGGTGCCTCATACTCCTCAGGTGCAGGGGCAGGGACAAGAGAAGGGGGAAGTA

    22346   TCATCCCAGCAGCCTCGCACACTCCCTAGGCTTCTACCCTCCCTCCTGATGTCCCTGGAA
            CAGGAACTCGCCTGACCCTGCTGCCACCTCCTGTGCACTTTGAGCAATGCCCCCTGGGAT
            CACCCCAGCCACAAGCCCTTCGAGGGCCCTATACCATGGCCCACCTTGGAGCAGAGAGCC
            AAGCATCTTCCCTGGGAAGTCTTTCTGGCCAAGTCTGGCCAGCCTGGCCCTGCAGGTCTC
            CCATGAAGGCCACCCCATGGTCTGATGGGCATGAAGCATCTCAGACTCCTTGGCAAAAAA
            [C,-]
            GGAGTCCGCAGGCCGCAGGTGTTGTGAAGACCACTCGTTCTGTGGTTGGGGTCCTGCAAG
            AAGGCCTCCTCAGCCCGGGGGCTATGGCCCTGACCCCAGCTCTCCACTCTGCTGTTAGAG
            TGGCAGCTCCGAGCTGGTTGTGGCACAGTAGCTGGGGAGACCTCAGCAGGGCTGCTCAGT
            GCCTGCCTCTGACAAAATTAAAGCATTGATGGCCTGTGGACCTGCTACAGTGGCCTGGTG
            CCTCATACTCCTCAGGTGCAGGGGCAGGGACAAGAGAAGGGGGAAGTAACCCCATCAGGG

    22476   ACAAGCCCTTCGAGGGCCCTATACCATGGCCCACCTTGGAGCAGAGAGCCAAGCATCTTC
            CCTGGGAAGTCTTTCTGGCCAAGTCTGGCCAGCCTGGCCCTGCAGGTCTCCCATGAAGGC
            CACCCCATGGTCTGATGGGCATGAAGCATCTCAGACTCCTTGGCAAAAAACGGAGTCCGC
            AGGCCGCAGGTGTTGTGAAGACCACTCGTTCTGTGGTTGGGGTCCTGCAAGAAGGCCTCC
            TCAGCCCGGGGGCTATGGCCCTGACCCCAGCTCTCCACTCTGCTGTTAGAGTGGCAGCTC
            [C,T]
            GAGCTGGTTGTGGCACAGTAGCTGGGGAGACCTCAGCAGGGCTGCTCAGTGCCTGCCTCT
            GACAAAATTAAAGCATTGATGGCCTGTGGACCTGCTACAGTGGCCTGGTGCCTCATACTC
            CTCAGGTGCAGGGGCAGGGACAAGAGAAGGGGGAAGTAACCCCATCAGGGAGGAGTGGAG
            GGTGCCTGAGCCGCCATGTGGGCATTGGGGGAGTGATGGGAATGCCAGCAGTGATGACGT
            TGACTACTGACTGAGCACCCACTACTATGACTGAGCACTCACTCGCTAGATACTATCTTG

    23143   GCCGGGCGTGGTGGCTCACGCCTGTAATCCCAGCACTTTAGGAGGCCAAGGCAGGTGGAT
            CACAAGGTCAGGAGTTTGAGATCAGCCTGGCCAACATGGTGAAACTCCATCTTTACTAAA
            AATACAAAAAATTAGCCAGGCATGGTGTTGCATGCCTGCATGCCTGTAATCCCAGTTACT
            TGGGAAGCTGAGGCAGGAGAATTGCTTGAACCCTGGAGGCGGAGGTTGTAGTGAGCCGAG
            ATCACGCCATTGCACTCCAGCTTGGGCAAGAAGAGAAACACTCTCAAAAAAAAAAAAAAA
```

FIGURE 3U

```
            [A,-,T]
            CAGGAAACTGGTGCTCAAAAAGGAAAAGTGACTCACCAAGGTCACAGACTAGGCAGTGAT
            GCTGGGGGAACCTGGCTCAGGGGACACAGACCTGGCCTGGGGCAGCCTTGCAGCTCCTCC
            ACTAAAATACTGAAAATGAGGGGCTTCGATGATGGTTATAATCGTATGGCAGAGCCCCAA
            CTCAACTGGAGCCCTGGGACCCAGAAGCTAGGGTCTCACTCCCTGCTTTTCCACAAGGCA
            CCATTAGGGCATCACCCCAGGCCTCGGCAGCCACGACGCAGGGATCCTGCCTCTCATTGG

23445       AGGAAACTGGTGCTCAAAAAGGAAAAGTGACTCACCAAGGTCACAGACTAGGCAGTGATG
            CTGGGGGAACCTGGCTCAGGGGACACAGACCTGGCCTGGGGCAGCCTTGCAGCTCCTCCA
            CTAAAATACTGAAAATGAGGGGCTTCGATGATGGTTATAATCGTATGGCAGAGCCCCAAC
            TCAACTGGAGCCCTGGGACCCAGAAGCTAGGGTCTCACTCCCTGCTTTTCCACAAGGCAC
            CATTAGGGCATCACCCCAGGCCTCGGCAGCCACGACGCAGGGATCCTGCCTCTCATTGGT
            [T,C]
            GGGGGCTTAGGGGCTCTGGGCTGCCCTCTTGAAGAGGGGGGTTCAGCCCAGCGAGGCACCC
            CCTATGCTGCACCCCACCAAGGTTAGGAAGAGGTCCTGTCCTCAGTGGGGCCCTCTGATG
            AACAGCCCATCAGGTCTGCGTCCACATGCCTTGGAAGAGATGGTGACATACTCAAAGTCC
            TTGAAGCCGCATATTAAACCACCTAGAGCACCATCTTCAAACATTTAGGGTCTGAGAAGA
            TAGGGGAAGTAAGCAATTTAAAACATTTCTTTATATTGGGCCAGGTGCAATGGCTCACGT

23974       GGTCTGAGAAGATAGGGGAAGTAAGCAATTTAAAACATTTCTTTATATTGGGCCAGGTGC
            AATGGCTCACGTCTGTAATCCCAGCGCTTTGGGAGGACGAGGATCACCTGAGGTCAGGAG
            TTCAAGATCAGCCTGGCCAACATGGAGAAACCCCATCTCTACTAAAAATACAAAAATTAG
            CTCAGGCGTGGTGATGTGCACCTGTAATCCTAGCTATTCAGGAGGCTGAGGCACAAGAAT
            TGCTTGAGTCAATATTGCACCACTGCACTCCAGCCTGGGCAACAGCGAGACTCTTGTCTC
            [A,-]
            AAAAAAAAAAAAGATATTTGCTGAAAAGACCCAGCCTGCCAAACTCAGGGGCAGCCAAGG
            GAGGTAGTGAAATGGAAGTTGGAGCTCAGCGCTCCCACACCTCCACTGCCCTCAGGCCTT
            CTCTGCCTCTTTCCCATCAGTCAGCTGCTTCTGGGCATGGTCCTGGCAGAGACTTGGCCT
            CCTTCCAGTTCAAGCTCCCTCTTAGATTGTGTCCCACGCCACTGAGTCTTTGGGACACTG
            GGTCAGATGTCTAGTCTGGCACAATTGGCAGGAATCCCAAGAAACAGTGTGAGTGAGGGG

24865       ATAGCTGGGATTACAGGTGTGCACCACCATGCCCAGCCTAATTTTTGTATTATTAGTAGA
            GATGGGGTTTCACCATGTTGTCCAGGCTGGTCATGAACTCCTGACCTCAAGTGATCCACC
            CGCTTTGGCCTCCCAAAGTGCTGGGATTACAAGCATGAGCCACAGTGCCTGGCCTGACCC
            TGCTCTTTTGAAAGACCATTCCCCCAAATTCTGTGCACCTGTGTGCCTTTCTTCTCTCTG
            CCTCCTCTCAGCTCTGCCCCGCTCTCCTCCCTTCTCCTCTGGCAAATCCCACTCATCTCT
            [G,T]
            GAAGCCCTTCTTCCAGGGGAAGCCCTGATCATGCTGCTTTCTCCTGTGGGAGGGATGAAG
            GACGTGGCCCACGGAGTTTGTTTTGTTTTGTTTTGAGATGGAGTTTTGCTCATGTTGCCC
            AGGCTGGGGTACAATGGTACGATCTCAGCTCACTGCAACCTCTACGTCCCGGGTTCAAGC
            GGTTCTCCTGCCTTAGCCTCCCCAGTAGCTGGGATTACTGGCATGAACCACCACACCTGG
            CTAATTTTGTGTTTTTAGTAGAGATGGGGTTTCTTCATGTTGGTCAGGCTGGTCTCGAAC

25040       GACCCTGCTCTTTTGAAAGACCATTCCCCCAAATTCTGTGCACCTGTGTGCCTTTCTTCT
            CTCTGCCTCCTCTCAGCTCTGCCCCGCTCTCCTCCCTTCTCCTCTGGCAAATCCCACTCA
            TCTCTTGAAGCCCTTCTTCCAGGGGAAGCCCTGATCATGCTGCTTTCTCCTGTGGGAGGG
            ATGAAGGACGTGGCCCACGGAGTTTGTTTTGTTTTGTTTTGAGATGGAGTTTTGCTCATG
            TTGCCCAGGCTGGGGTACAATGGTACGATCTCAGCTCACTGCAACCTCTACGTCCCGGGT
            [T,C]
            CAAGCGGTTCTCCTGCCTTAGCCTCCCCAGTAGCTGGGATTACTGGCATGAACCACCACA
            CCTGGCTAATTTTGTGTTTTTAGTAGAGATGGGGTTTCTTCATGTTGGTCAGGCTGGTCT
            CGAACTCCCAACCTCAGGTGATCTGCCTGCCTCGGCCTCCCAAAGTACTGGGATTACAGG
            GTTGAGCCACTGTGCCTGGCCCAGGCCCACGGAGTTTTAAGAGGCTTCCTGTGGCAGTGG
            CATCCAGACGGAGTGCAGAAACTCAAAGTTGAAGGCCAGAAGCTCAGGGAAGGGGGAGTG

25046       GCTCTTTTGAAAGACCATTCCCCCAAATTCTGTGCACCTGTGTGCCTTTCTTCTCTCTGC
            CTCCTCTCAGCTCTGCCCCGCTCTCCTCCCTTCTCCTCTGGCAAATCCCACTCATCTCTT
            GAAGCCCTTCTTCCAGGGGAAGCCCTGATCATGCTGCTTTCTCCTGTGGGAGGGATGAAG
            GACGTGGCCCACGGAGTTTGTTTTGTTTTGTTTTGAGATGGAGTTTTGCTCATGTTGCCC
            AGGCTGGGGTACAATGGTACGATCTCAGCTCACTGCAACCTCTACGTCCCGGGTTCAAGC
            [G,A]
            GTTCTCCTGCCTTAGCCTCCCCAGTAGCTGGGATTACTGGCATGAACCACCACACCTGGC
            TAATTTTGTGTTTTTAGTAGAGATGGGGTTTCTTCATGTTGGTCAGGCTGGTCTCGAACT
            CCCAACCTCAGGTGATCTGCCTGCCTCGGCCTCCCAAAGTACTGGGATTACAGGGTTGAG
            CCACTGTGCCTGGCCCAGGCCCACGGAGTTTTAAGAGGCTTCCTGTGGCAGTGGCATCCA
            GACGGAGTGCAGAAACTCAAAGTTGAAGGCCAGAAGCTCAGGGAAGGGGGAGTGTGAGTT

25193       ATCATGCTGCTTTCTCCTGTGGGAGGGATGAAGGACGTGGCCCACGGAGTTTGTTTTGTT
            TTGTTTTGAGATGGAGTTTTGCTCATGTTGCCCAGGCTGGGGTACAATGGTACGATCTCA
```

FIGURE 3V

```
        GCTCACTGCAACCTCTACGTCCCGGGTTCAAGCGGTTCTCCTGCCTTAGCCTCCCCAGTA
        GCTGGGATTACTGGCATGAACCACCACACCTGGCTAATTTTGTGTTTTTAGTAGAGATGG
        GGTTTCTTCATGTTGGTCAGGCTGGTCTCGAACTCCCAACCTCAGGTGATCTGCCTGCCT
        [C,T]
        GGCCTCCCAAAGTACTGGGATTACAGGGTTGAGCCACTGTGCCTGGCCCAGGCCCACGGA
        GTTTTAAGAGGCTTCCTGTGGCAGTGGCATCCAGACGGAGTGCAGAAACTCAAAGTTGAA
        GGCCAGAAGCTCAGGGAAGGGGGAGTGTGAGTTGAGGAGTCTCTTGGCTGCCAGGGCCAG
        AAACCGAACTCCAAGCCTCTCCACAACAGCGGGTGTAGAGCATGTAGAATCAGAGAGGAG
        GCTGAGCCATGCAGCCCCGAGAAGAGGGGAATGCCACTGAGCCACAGAGACCCAGTGCCA

25529   AGTGCAGAAACTCAAAGTTGAAGGCCAGAAGCTCAGGGAAGGGGGAGTGTGAGTTGAGGA
        GTCTCTTGGCTGCCAGGGCCAGAAACCGAACTCCAAGCCTCTCCACAACAGCGGGTGTAG
        AGCATGTAGAATCAGAGAGGAGGCTGAGCCATGCAGCCCCGAGAAGAGGGGAATGCCACT
        GAGCCACAGAGACCCAGTGCCACTGCCAGGTGTCTCTGCCTCCACTTCCCATGACCC
        [T,G]
        GCCTGTCTCTGTATGCAGGCTTCACCCTCTCTCGTTGTACATTGTACACATTCTAGGTGA
        CACCAGCAGCTTCTGATTCTCATCTCCCATAACATCAGCCCCCCAGAGAGGGGACAACTG
        CTGAGCTGATAACATAATAGATGCCCCTTTCCTGGAGGCCATGGTCATGGTCAGCGTGGA
        GAGGATGAAGCCTGAGCAGGCAGGATCGGGGGTCTAGAGGGGAAGGAGGTGGAAGTT

25614   GGCCAGAAGCTCAGGGAAGGGGGAGTGTGAGTTGAGGAGTCTCTTGGCTGCCAGGGCCAG
        AAACCGAACTCCAAGCCTCTCCACAACAGCGGGTGTAGAGCATGTAGAATCAGAGAGGAG
        GCTGAGCCATGCAGCCCCGAGAAGAGGGGAATGCCACTGAGCCACAGAGACCCAGTGCCA
        CTGCCAGGTGTCTCTGCCTCCACTTCCCATGACCCGGCCTGTCTCTGTATGCAGGCTTCA
        CCCTCTCTCGTTGTACATTGTACACATTCTAGGTGACACCAGCAGCTTCTGATTCTCATC
        [C,T]
        CCCATAACATCAGCCCCCCAGAGAGGGGACAACTGCTGAGCTGATAACATAATAGATGCC
        CCTTTCCTGGAGGCCATGGTCATGGTCAGCGTGGAGAGGATGAAGCCTGAGCAGGCAGGA
        TCGGGGGTCTAGAGGGGAAGGAGGTGGAAGTTGAGATCACAGACCTGTGGTCAGGTGGCC
        TGGGAAGGGTTTGACGAGTGTCGGCCCAAAGAGCTTGGAAGGGATTTTGCTGCTGTGGGT
        GAGCACTGCCTCTCCCCTTAGGGACAACAGCCACCTCTTCTCTCCCCATTTGCCTTTCCC

25785   CCAGTGCCACTGCCAGGTGTCTCTGCCTCCACTTCCCATGACCCGGCCTGTCTCTGTATG
        CAGGCTTCACCCTCTCTCGTTGTACATTGTACACATTCTAGGTGACACCAGCAGCTTCTG
        ATTCTCATCTCCCATAACATCAGCCCCCCAGAGAGGGGACAACTGCTGAGCTGATAACAT
        AATAGATGCCCCTTTCCTGGAGGCCATGGTCATGGTCAGCGTGGAGAGGATGAAGCCTGA
        GCAGGCAGGATCGGGGGTCTAGAGGGGAAGGAGGTGGAAGTTGAGATCACAGACCTGTGG
        [C,T]
        CAGGTGGCCTGGGAAGGGTTTGACGAGTGTCGGCCCAAAGAGCTTGGAAGGGATTTTGCT
        GCTGTGGGTGAGCACTGCCTCTCCCCTTAGGGACAACAGCCACCTCTTCTCTCCCCATTT
        GCCTTTCCCTTCTGTAGATATGAAACACAGGCCTCCTTGTCAGGCCCCTACTTAACCTCC
        GTGATGGGGAAAGCGGCCGGAGAAAGGGAGTTCTGGCAGCTTCTCCGAGACCCCAACACC
        CCACTGTTGCAAGGTGAGTCATGGCCTGACACTCTGGATGTGTCCCCTACCCCAAGCTTA

26266   GTGATGGGGAAAGCGGCCGGAGAAAGGGAGTTCTGGCAGCTTCTCCGAGACCCCAACACC
        CCACTGTTGCAAGGTGAGTCATGGCCTGACACTCTGGATGTGTCCCCTACCCCAAGCTTA
        CTCAGCCAAGAGGCTTCATCAACTCACCCCAGCTTTCCCTAGCACCCTCCTGGGCCACAC
        CTTCACAAAATCACTGATGCTCAAAGTTGGATATAATATATTGAACTGAAGCCTTAGCAT
        TTTTATGCAAGTTACTGTGGAAATTCTAGGAAACCAGACAGATTACAAAAAAAAAAAAAA
        [A,-]
        CTAGAAGAAAATTAACATCACCTAGGATATACTACCTAGGAATAACGTCTTTTATTTTGA
        GATGGAGTTTCGCTCTTGTTGCCCAGGCTGGAGTGCAGCGGTATGATCTCGGCTCGCTGC
        AACCTCCGCCTCCTGGGTTCATGTGATTCTTCCACCTCGGCCTTCCTAGAGCCCAAGTGG
        TCTGCCTGCCTCTGCCTCCCAAAGTTCTGGGATTACAGGCATGAGCCACCGCACCCAGCC
        AAAATTACTTAACTTTTCTTCTAGATACTTTTTAAAAAATATGGCAGTAAGTTTTTCATAA
```

FIGURE 3W